

Unesco's crisis is for all of us

The US withdrawal from Unesco is intelligible and may be wise. The remaining members should abandon gentility in the hope of winning overdue reform. Unesco alone is not the worry.

DISMAY about the proposed withdrawal of the United States from the United Nations Educational, Scientific and Cultural Organization (Unesco) is natural. Especially when the US Administration is more than usually insensitive to the needs of collaboration overseas, the prospect that the United States may withdraw from a surviving network of international collaboration must be a matter for regret. But those in Washington who have been wringing their hands about the proposed withdrawal (see p.98), officers and officials of the National Science Foundation and the National Academy of Sciences, should not go too far, or they will find themselves in the same case as those, in the 1930s, who applauded Mussolini because he had arranged that Italian trains should run on time. The fact is that the US withdrawal from Unesco has been decided for reasons that have nothing to do with the organization's scientific programme and that, if valid, not merely justify the withdrawal but make it necessary.

For the past three years, the central complaint against Unesco has been its espousal of what it calls the "new world information and communications order", NWICO for short. This notion, which has been rumbling about in Le Corbusier's building in Paris since 1972, is a vivid illustration of the mischief that can be done when people set out to translate good but ill-defined intentions into rules that the whole world is then expected to follow. The argument has gone like this. Free speech is a common good, right? *Of course*. So everybody should be free to communicate, right? *Well . . . yes*. But it's not like that, because the world's successful news agencies and television networks are mostly owned by capitalists and put out distorted information about countries against which they are prejudiced but which lack the means to answer back; so should there not be some way in which governments not in control of international news agencies or television networks can influence what is said about them by others, and in the meantime should not journalists operating internationally be licensed by those about whom they would write? *Well . . . no*.

Doublespeak

Unesco's remedy would be worse than the disease it is meant to cure. To regard as synonymous the principle of free speech for people and the doctrine of information control by governments is, in 1984 language, doublespeak. To nurse the ambition, as Unesco has done for the past few years, to embody this mischievous obfuscation in some kind of international convention would run flatly against its own objectives of helping to make the world a more civil place. Unesco's management has been obtusely unwilling or unable to see the difficulty, citing the acknowledged imperfections of news agencies, newspapers and television networks as proof that something should be done. It has steadfastly refused to understand that many of its member governments, but especially that of the United States, are constitutionally bound to the view that the imperfections of the international press, in any case remediable by means other than NWICO, are the price that must be paid for free speech. To general surprise, the management seems to have soft-pedalled on the issue at Unesco's General Conference in Paris last November, even accepting a British gloss on the description of NWICO as an "evolving and continuous process", which is apparently a diplomatic device for making sure that the new order cannot now be brought in by fiat. Yet the issue has not gone away, and those

remaining within Unesco will be battling with it for years to come.

If the State Department's calculation were that pulling out of Unesco is the best way of making sure that this tedious matter does not rear its head again, or that the threat of doing so may be the most effective way of changing Unesco's management, the loss to the scientific community would be a small price to pay — as would have been unpunctuality on the Italian railways in the 1930s for some other kind of political leadership than that provided by Mussolini.

The weakness in the position the United States has taken is threefold. Resignation would have been more convincing at the end of 1982, when the row about NWICO was at its height, than now, when the issue seems to have gone off the boil. In present circumstances, it is hard to tell to what extent the decision reflects the general impatience with United Nations institutions that keeps bubbling up within the administration. And there is bound to be a suspicion that resignation from Unesco, or the threat thereof, is an early move in November's presidential election campaign. If it wishes to avoid the shabby conclusions to which these suspicions point, the State Department should quickly give substance to the hints it has been dropping in Washington and put up the money that will be needed before the end of the year if the important parts of Unesco's programme are not to founder. The need that the International Council of Scientific Unions (ICSU) should not be endangered typifies what must be done.

Disenchantment

There is also an urgent need that the State Department, and the US Administration as a whole, should come to a clearer and more guileful understanding of the reasons why its relations with UN organizations have become so irksome. One is that the United States has allowed its discovery in the 1940s that the United Nations did not obviate the need for international organizations, its alarm at the total cost (to which it is the chief contributor) and its dislike of being bested by the political manoeuvrings of its critics, to engender a disenchantment and a detachment that only exaggerates legitimate discontents. The United States would be better advised to define what is wrong and to seek to put it right.

There are general complaints against the UN system as a whole, and particular complaints against specific UN agencies, all of which unhappily apply to luckless Unesco. First, the agencies are needlessly bureaucratic, with decisions on minor matters rubber-stamped (or sometimes capriciously overturned) by officials in a complicated hierarchy whose exaltation and competence are often imperfectly correlated. Second, and not simply as a consequence, many of the agencies are outrageously expensive ways of accomplishing inexpensive tasks. Third, the agencies are often disastrously politicized, which means that issues entirely irrelevant to the objectives of an agency (such as whether Israel should continue as a member of the World Health Organization, a bone of contention three years ago) dominate proceedings. In this connection, it is only proper to acknowledge that Unesco's concern with the international traffic in information is not irrelevant to its terms of reference, although a majority of Unesco's member governments are guilty of imprudence for having used their voting power to advance the issue, just as it was shabby of them to use the same voting power to lift the minimum number of a member's nationals whom Unesco must employ from four to five.



All of these complaints are legitimate, but can be sensibly assessed only against the knowledge that the UN system and its agencies are the only systematic means by which nations with very different interests are regularly required to meet each other. In recent years, the process has done very little to sustain the hope of the United Nations' founders that the system would somehow abolish conflicts large and small, but it has become a powerful channel for the flow of multilateral technical and material assistance from developed to developing countries, itself an indispensable complement to what countries are able to do bilaterally. Latterly, the UN system has become indispensable (if infuriating) as the only systematic meeting ground between the two kinds of governments. The value of all this is not negligible even if the present cost is too high.

Moreover, the complaints against the United Nations and its agencies are not irremediable, as the performance of the more efficient agencies shows clearly enough. The World Meteorological Organization is top of most people's lists of good performers, the International Atomic Energy Agency plays an indispensable part in the administration of the Non-Proliferation Treaty, the International Telecommunications Union is indispensable (if too much a prisoner of national telecommunications monopolies and thus in need of reform), the Tropical Diseases Programme of the World Health Organization has surprised the sceptics by having emerged as a sensibly directed programme of research and development and not the slush fund that seemed in prospect ten years ago. Arranging that Unesco should function that well is bound to be difficult, given that this agency's terms of reference are more fuzzy than any other's, but there is much that can and should be done to simplify Unesco's housekeeping.

The administrative problems should be more easily soluble than most Unesco delegates will allow. The bureaucracy and high cost, consequences though they are of the way in which the United Nations and its agencies are instruments of the collective will of member governments, should with wit and perseverance be remediable. Unesco as it has become is two things — a talking shop and an executive agency. It is inevitable and even proper, within the agency's terms of reference, that all member governments should feel free to have their say on grand matters of policy, but intolerable and unnecessary that the same should apply to the selection of people who should benefit from training programmes in special fields, or to the management of collaborative scientific programmes of the several kinds on Unesco's books. For an agency whose ambitions forever outstrip resources and even common sense, but whose staff is bloated by the small army of professional people hired to second-guess hostile governmental delegations, the sensible and frugal course would be to delegate much of this work. The International Council of Scientific Unions, already a powerful source of help to Unesco (but which antedates it) could for example take over much of what is intended. Then at least there would be some hope that the overhead cost, which often amounts to a third or more of what Unesco has to spend on its programmes, would be reduced.

Puzzlement

Many other Unesco projects should, on the other hand, be promptly brought to an end. Even in the relatively tangible field of science and technology, it is hard to see why Unesco is planning to spend \$85,000 this year on helping member states to draw up "national programmes for research in biotechnology and its applications", and hard to imagine what special expertise it pretends to have in this competitive field. It is similarly mystifying that it should wish to spend \$89,000 to help national governments to work out strategies for the development of information technology (or "informatics" in Unesco language). If there were even the most meagre evidence that member governments were anything but over-aware of the opportunities in these fields, the modest sums involved might be forgiven, but as things are they will literally be wasted. Unesco, unfortunately, has never been able to exercise self-restraint but instead feels bound to jump on every new bandwagon that comes along, usually with too small an

amount of money to be decisive, often with too woolly an idea of what might be done to be helpful and sometimes with the ill-concealed intention of stealing other people's thunder so as to be able to magnify its own appearance on the scene.

But how can an incorrigibly expansionist organization such as Unesco be made to behave more sensibly? Those member governments dismayed by such developments must not underestimate their power. While the massed votes of the developing countries in the group of 77 (misnamed because there are many more of them) can almost always force the general conference to accept unpalatable proposals, particular wrongheadedness can be rooted out whenever delegations dig in their heels, abandon the customary gentilities of UN meetings which require that the chairman of every committee should be applauded for his eminence and congratulated for his wisdom, throw out the dangerous notion that a nonsense will become sensible if enough people vote for it, and simply declare that they will not be parties to proposals that they know to be objectionable. That, of course, would demand more of delegations than they have in the past been able to provide. The present crisis springs from members' past neglect.

Politicization

The politicization of Unesco is a more difficult problem, although not very different in its essentials. The first need is that all member governments should recognize that an organization with such wide terms of reference cannot be apolitical. If Unesco's programme on the free flow of information had not turned out to offend North America and Western Europe, it would no doubt have caused similar offence to the governments that practise censorship. Yet who will say that the free flow of information is not relevant to the interests of a body with science, education and culture in its title? The now-complainant governments have weakened their own positions in the past by failing to anticipate that such conflicts are unavoidable, to argue with conviction that there are two sides to most of the coins turned up — and then to insist on what is plain for all to see, that neither Unesco nor any other such representative body is equipped to be a negotiating forum for the removal of irreconcilable differences.

The second need is somehow to neutralize the irresponsible use of the votes of the group of 77, both in Unesco and the other UN agencies. Here again, the complainant governments are not nearly as powerless as they think. If, as this year, Unesco is determined to hold a major conference on the future development of information technology of which those with the most to contribute disapprove, they have a simple remedy — simply not to participate. If the massed votes should commit Unesco to a resolution calling for the restoration of all cultural property to its place of origin, or to a demand that intellectual property should be freely available to some class among the membership, they can do more than merely decline to cooperate — they can, more openly than has been their habit, argue for the notion that equity is not unidirectional. The complainant governments have in the past been too unreflective in their response to the demand that all countries should at all times be qualitatively equal, when the best that can be hoped for is dynamic diversity. Plain speaking is again the answer.

The future for Unesco is in all the circumstances not nearly as gloomy as the threatened withdrawal of 27 per cent of its budget would suggest. The ideal is that the hard times ahead should persuade the organization and its member governments that three-quarters of the budget is ample to support the present programme. Only last week, the director-general was asking his staff to reduce spending by 10 per cent this year — nearly half the battle. The complainant governments not planning to withdraw know what they have to do to encourage him in this direction — be less polite and better-informed. The United States should plan seriously to practise what Mr George Schulz promised in the State Department's letter of resignation — that the United States will collaborate with Unesco in (and help pay for?) those programmes it approves of. The important need is that disenchantment with Unesco should not infect US policy to the rest of the UN system. □

Monoclonal patent disputed

Disclosure threat to Japanese rights

Washington

THE Wistar Institute's efforts to patent monoclonal antibodies have run into trouble in Japan and the United Kingdom. Twenty-seven separate challenges have been filed in Japan to a Wistar patent for monoclonals that specifically attack tumour cells. The patent is based on research by Hilary Koprowski and Carlo Croce of the Wistar Institute and was assigned to Wistar by the researchers. And the British Patent Court recently affirmed the rejection by the Patent Office there of a second patent that covers monoclonals specific for viral antigens. Both have been successfully patented in the United States.

The challenge to the tumour antibody patent in Japan may prove the more costly of these setbacks to Wistar. A test kit that makes use of the antibodies for diagnosing solid tumours is already on the market in Japan, and Wistar has already sold three licences to the US patent. Clinical trials of anticancer agents that use the antibodies to home in on tumour cells are under way in the United States. Although Wistar's associate director, Warren Cheston, says that "realistically" he expects the patent to bring in less than \$100,000 a year to the institute, he acknowledges that "if you believe the people on the commercial side, there's potentially huge sales out there". Cheston points out that the non-monoclonal diagnostic for colo-rectal cancer now in use is a multi-million dollar market.

Those challenging the patent are basing their case largely on an article that appeared in *Nature* the day before the US patent application was filed. The US patent is not threatened by the Japanese challenge, but the US filing established the priority date of the invention. And unlike the United States, which allows an inventor to apply for a patent up to a year after disclosing his invention, Japan (like Western Europe) demands "absolute novelty" — no disclosure before the first filing. The patent was filed for in the United States on 28 April 1978; in the 27 April issue of *Nature*, a News and Views article appeared with a brief description of Koprowski's work.

According to Helmuth Wegner, a Washington patent attorney who is representing one of the challengers, that disclosure, even if it is no more than "a fairly broad suggestion", is enough to invalidate the Japanese patent. The *Nature* article provides no experimental details; it notes in a single sentence that Koprowski had produced hybridomas that secrete antibodies with tumour-specific activity. The legal issue is whether that disclosure is in theory sufficient to allow someone else to

duplicate Koprowski's work with a reasonable amount of experimentation.

Another tricky issue is sure to be the exact date that the *Nature* disclosure is considered to have taken place. The magazine would have been available in *Nature*'s London office on 26 April, and would have been posted to subscribers on that day. The Japanese law defines disclosure to have occurred when the publication is "distributed", a term not yet tested in court.

The report referred to (*Nature* 272, 751; 1978) is by Dr Elizabeth Simpson of the Medical Research Council's National Institute of Medical Research at Mill Hill, London, and describes the proceedings of a workshop organized by M. Potter and F. Melchers at the National Institutes of Health on 3–5 April 1978. Simpson refers to the production of monoclonal antibodies against "Differentiation and tumour antigens" by "several groups", and says that Koprowski, using human cancer tissues as a source of antigens, had

been able to produce monoclonal antibodies which reacted with human cells so as to "suggest that some of them may be against tumour-specific antigens".

If the challenges are upheld, Wistar would lose the Japanese market, which is the second largest pharmaceutical market in the world. Patents in North America and West Germany would not be affected.

The British patent covered the work by Koprowski, Croce and Walter Gerhard on monoclonal antibodies specific for viral antigens. The Patent Office had earlier ruled that the work was an obvious application of the basic procedures for producing monoclonal antibodies developed several years earlier but never patented by César Milstein and Georges Kohler of the Medical Research Council's Laboratory of Molecular Biology. The Patent Office noted in particular an article that had appeared in *The Lancet* — before the filing of the Wistar patent — that suggested a number of possible uses for monoclonal antibodies, including the detection of viral antigens. Wistar, which had little difficulty obtaining broad patent protection in the United States and several European countries for the viral antibodies, appealed against the rejection; late last year the patent court ruled in favour of the Patent Office.

Stephen Budiansky

European forests

Pollution, pathogens and pests

FORESTS in the Federal Republic of Germany (FRG) and many other parts of Central Europe are deteriorating rapidly. It is estimated that in West Germany 34 per cent of all forest, which covers 27 per cent of the country's total land area, is already damaged. The city of Augsburg has created a special nature park to inform its citizens about the advancing disaster and the Christmas tree business has languished. Whereas in 1982 damage to conifers was estimated in Bavaria at 10–30 per cent, the official figure is now 60 per cent for Scots pine (*Pinus sylvestris*) and 80 per cent for fir (*Abies*). Furthermore, 44 per cent of the beech forest is showing signs of distress.

The situation in Switzerland, where the first symptoms were reported at the Federal Institute of Forest Science in early 1983, has deteriorated rapidly, exacerbated by heavy snowfalls and storms in the two previous winters. The first signs of distress were shown by *Abies alba* in the eastern and northern plateau that includes Berne, the Bodensee and the most heavily populated areas. Three months ago, reports came in from the southern Alps of damage to spruce (*Picea excelsa*), the dominant species, between 1,000 and 1,800 m altitude and in November of massive damage in some localities.

In Czechoslovakia, 11 per cent of the forest is said to be destroyed and a further 11 per cent dying. A recent ecological report by the Czech Academy of Sciences,

which was not published in Czechoslovakia but a copy of which reached the French newspaper *Le Monde* through the human rights organization Entraide et Action, draws a grim environmental picture that forecasts that 45–60 per cent of the nation's forest will be degraded by the end of the century.

Although most of the damage is caused by "acid rain", two serious forest pests are adding to the toll. In Switzerland, winter weather damage was followed by an exceptionally hot summer in 1983. Attacks by the bark beetle *Ips typographus* L. reached serious levels. This beetle normally produces two generations per summer, but in 1983 it produced three. This winter so far has been exceptionally mild and the beetle will overwinter in serious numbers if the normal cold fails to occur. In the stressed environments at high altitudes, regeneration of forest is slow and the Swiss institute is seriously concerned that the threatening deafforestation will lead to erosion and avalanches.

In Poland, the moth *Lymanthia monacha* (the black arches moth) known in German as *Die Nonne* (the nun) is said to be the major cause of damage. Historical accounts exist of this moth occurring in devastating swarms from Europe through Central Asia to China and Japan, killing large areas of both conifer and broad-leaved forest, until it eventually succumbs to its own disease.

Sarah Tooze

French research support

More clout for CNRS

FRENCH university scientists should be bracing themselves for a radical new year. Last week, the Council of Ministers, the French cabinet, considered the programme of the Centre National de la Recherche Scientifique (CNRS), the body that supports most basic science in France, and has publicly approved of it.

Why should the council do that? To give CNRS political weight in dealing with ministries other than the ministry of research and industry, to which it belongs, says Pierre Papon, CNRS director-general. Among those ministries is the ministry of national education (MEN), responsible for the universities.

CNRS will be restructuring its relationship with the universities in 1984, says

Papon. The 75 universities of France are "too many", which has led to a "balkanization" of university politics. According to Papon, CNRS will in future no longer be able to support all subjects in all universities. All this is further evidence of a tightening grip by CNRS of the distribution of its money in the university system and of an increasing emphasis on its own laboratories. But the picture is not all black for university researchers supported by CNRS. Papon insists that the total value of support will not fall — it will just be distributed differently.

In the past, CNRS guidelines for university support have centred on "excellence". Exactly what the research was about was not of great concern. Now, excellence will be necessary but not sufficient, says Papon; "coherence" with CNRS policy will also be expected. This year, for example, those policies include a regional dimension — a wish to improve support in the north and west, relatively undeveloped scientifically; an emphasis on interdisciplinary research, the barriers between disciplines being felt to be particularly strong in France; and special support for life sciences.

Clearly, this marks an increase in "directive" rather than "responsive" support for science, a matter of debate in many other countries. But France is to go still further. CNRS is to sign — possibly this week — an outline agreement with the "Direction de la Recherche", which with MEN is responsible for the universities. This will mark a new degree of cooperation between the two bodies. Some will fear this development as a device for extending CNRS policy now coloured strongly by its exposure in the ministry of industry and research, and by the aim of the present government of achieving economic growth through science and technology — into university research.

Both MEN and CNRS insist, however, that there will be room for independent policies. There is, however, considerable unease in the universities about the way in which CNRS support and influence will develop, particularly among those who do not see their work as of particular economic relevance. And it has been exacerbated by the removal of "individual aids" — a system whereby CNRS supported individuals with two or three technicians, as a step to becoming an "associated laboratory" of CNRS shared between university and the centre. There are not enough technicians to go round, says CNRS. So such early support is to be limited to a few selected "young groups", while individuals will apply for "free contracts" — two-year grants without technicians that are not designed to lead to associated status.

Robert Walgate

Unesco

US scientists hit by withdrawal

Washington

THE United States' decision to withdraw from the United Nations Educational, Scientific and Cultural Organization (Unesco) by the end of this year has dismayed many US scientists who have worked closely with the agency. They say that Unesco's scientific programmes have generally remained free from the political posturing that has tainted the agency's other activities, and warn that withdrawal may exclude US scientists from valuable international data-sharing agreements.

Formal notice of the decision was delivered to Unesco at the end of December. Under United Nations regulations, the withdrawal will not take effect until 31 December 1984, and the United States will therefore continue to pay its \$50 million contribution (about a quarter of Unesco's budget) for the remainder of this year. Explaining the American decision to pull out, a State Department spokesman said Unesco had politicized virtually every activity it conducted, mismanaged its budget and routinely adopted policies hostile to the institutions of a free society.

Both the National Academy of Sciences and the National Science Foundation (NSF) told the State Department at the end of last year that a decision to pull out could have damaging repercussions for US science. A letter from the academy characterized science as one of Unesco's success stories and pointed out that no alternative machinery existed to service international collaboration concerned with such subjects as the oceans, climate, the solid Earth and the biosphere.

NSF has made a detailed study of Unesco's scientific value for the State Department but its findings have not yet been published. The general view of scientists with knowledge of Unesco, however, is that the benefits of staying in outweigh those of leaving:

- Dr Della Laura, chief of international hydrology at the US Geological Survey, says Unesco's International Hydrology Programme has an excellent record and has successfully steered clear of political issues. If US scientists are unable to continue their participation, she says, the United States will "greatly miss" an arena in which global information of hydrological problems is rapidly exchanged.

- Unesco's activities in the Earth sciences have received similar praise from Dr Lin Hoover, deputy chief of international scientific programmes at the Geological Survey. About 300 American geologists a year take part in research under the Unesco-sponsored International Geological Correlation Programme and their exclusion will force the United States to negotiate countless bilateral agreements

Concerted selectivity

IN its unusual declaration, the French cabinet has offered its support to CNRS in pursuing a policy of "selectivity and concentration", which echoes the policy of a certain Science and Engineering Research Council across the Channel. In outline, the declaration announces that CNRS will increase cooperation with "all involved in research", including:

- The universities, with which CNRS will sign individual and collective agreements. The declaration foresees "decisions, which will allow the creation of new associations [joint research groups, the principal form of CNRS support] or the closure of others, taken in liaison with the universities and the minister of education, and taking into account the quality of the research, its relevance to national scientific priorities and the balance between regions".

- Other research councils. The declaration singles out INRIA (informatics), CNET (telecommunications), and CNEXO (ocean science).

- Industry, where the declaration foresees more collaboration agreements between CNRS and industry.

- "Social partners", a neologism for trade unions, involving increased research on the impact of new technologies.

- And the regions ("states"), which will see more conventions lasting several years such as those signed with Nord-Pas-de-Calais and Provence-Côte d'Azur, relating CNRS policy to local needs perceived by the (political) regional councils.

The declaration also states that "voluntary mobility of researchers will be encouraged". And this could indeed be an important step for CNRS, for it would allow, for example, a CNRS scientist to leave his or her post for up to 3 years to work in industry, freeing that person's salary and position for the period. This could dramatically increase the fluidity of CNRS laboratories.

Robert Walgate

and will reduce opportunities for US geologists to visit foreign sites in the company of international experts.

Dr Roger Revelle, professor of science and public policy at the University of California, San Diego, says the United States has been deeply involved in Unesco's Intergovernmental Oceanographic Commission and two major projects it is helping to coordinate — Tropical Oceans and Global Atmosphere (TOGA) and the World Ocean Circulation Experiment (WOCE). He says the intergovernmental machinery established by Unesco is essential to gain permission to set up tide gauges and other equipment in territorial waters.

The State Department has hinted during informal conversations with the scientific community that some of the money now given to Unesco could be diverted to the apolitical and, some believe, more efficient International Council of Scientific Unions (ICSU). But the suggestion has been

dismissed by the National Academy of Sciences, which insists that only a governmental-level organization can command the prestige and resources necessary to co-ordinate global research programmes.

The academy is not, however, wholly uncritical of Unesco. Its letter to the State Department acknowledges that its management of scientific programmes could be improved. Now that the United States has announced its firm intention to leave, US scientists who work with Unesco hope that they will nevertheless be allowed to participate in individual programmes of obvious benefit to the United States. The United States intends, for example, to remain within the copyright convention and the International Programme for the Development of Communication. The State Department has left the door open, but it remains to be seen whether Unesco itself, jilted by its biggest paymaster, will agree.

Peter David

Solar energy

Sahara's power for Europe?

Rehovot

A SCHEME that may one day permit Sahara sunshine to power the factories of Europe was discussed this week by participants in an International Workshop at the Weizmann Institute of Science in Rehovot and at Ein Bokek on the shores of the Dead Sea.

The idea is to collect solar energy in a desert area (say Israel's Negev, the south-western part of the United States or the Sahara) and to transform it into energy-rich chemicals for piping to industrial areas further to the north where, in yet another catalytic transformation, this energy would be released as heat for use in manufacturing processes.

Professor Israel Dostrovsky of the Weizmann Institute, organizer of the workshop, points out that "solar heat has generally been used close to the site where it is collected. This is because transforming it into electricity and then back to heat (if that is what the customer needs) entails a loss of about 75 per cent of the energy." He hopes, however, that the "thermochemical pipeline" will make it possible to deliver solar heat to a remote customer with much smaller losses and, simultaneously, provide a convenient way to store solar energy for night-time use.

Participating in the workshop were US and West German scientists who have studied the transformation of surplus nuclear energy into chemical energy with much the same goals in mind. Indeed, the Germans have even built a successful 10-megawatt pilot plant at Jülich, near Cologne. But solar energy presents an additional challenge because, unlike nuclear energy, it does not come in a steady flow.

A one-megawatt pilot plant to test the "thermochemical pipeline" concept will be built on the Weizmann campus with

energy from a solar tower.

While institute researchers have no delusions about solar sources supplying most of the world's energy needs, they think the new scheme is worthy of support — and particularly from the oil sheikhs of Saudia Arabia and the Gulf States. "If we're successful", Professor Dostrovsky observes, "their vast stretches of sunbaked sand will still be profitable after their oil wells run dry."

Another original Israeli development in the solar sphere won international recognition two weeks ago when a contract was signed between the Edison Company in Southern California and Ormat Turbines

of Yavne, south of Tel Aviv. The Israeli firm will be responsible for the construction and operation of a solar-pond power plant, expected to supply 12,000 kilowatts of power to California Edison customers by the end of 1985, and 48,000 kilowatts in 1987 (with the addition of three more ponds).

Ormat operates two solar ponds at the Dead Sea; the first, inaugurated in 1979, produces 150 kilowatts, and the second, opened last year, produces 5 megawatts. They exploit the fact that in a standing pond of salt water a salinity gradient is created, one which they have learned to maintain artificially. Temperatures at the very salty bottom reach up to 90°C, not very hot compared with the 500°C steam used in conventional turbines but hot enough to drive the special Ormat turbine generators in which fluid with a low boiling point is substituted for steam.

Here, as elsewhere, scientists and politicians now speak frequently about what they expect to be happening by the year 2000. By then, Professor Dostrovsky hopes that his "solar pipeline" will be contributing significantly to a solution of energy problems, and the Israeli Ministry of Energy predicts that 11 per cent of the country's total energy needs will be met from solar sources, divided in the following way: 12 per cent from "passive" energy (direct absorption not using collectors), 30 per cent from flat-plate collectors, 30 per cent from concentrating collectors producing industrial process heat, 16 per cent from photovoltaic systems producing electricity directly, and 12 per cent from solar ponds.

In another 16 years, it will be possible to know whether this extremely optimistic prediction proves to be correct.

Nechemia Meyers

A place in the Sun

THE multitude of solar energy projects in Israel, were reviewed in a report recently presented to the Food and Agricultural Organization of the United Nations by Israel's Agricultural Research Organization. The authors, D. Groves and I. Segal, point out that more than 700,000 Israeli households now have their water heated by solar energy, thus accounting for 65 per cent of all domestic water heating. The result they say, is a saving of 6 per cent in electricity generation and 2 per cent in all primary energy supply. "These figures", they add, "make Israel the world's largest *per capita* solar energy user."

Almost all this water is heated by flat-plate collectors, developed in the 1950s by Dr Harry Tabor and now manufactured by no fewer than 130 individual companies.

Some of Israel's most interesting solar energy work is being done at the Desert Research Institute of Ben-Gurion University of the Negev in Sdeh Boker, the isolated place at which solar enthusiast Ben-

Gurion made his home after retiring from the premiership. At Sdeh Boker, one finds, for example, a special adobe house — constructed by Michael Kaplan and Eli Levin-Epshtein — which requires only US\$12 for total winter heating as against an average of \$1,000 for neighbouring homes.

Solar energy is also of interest to Israeli farmers. It is, for example, being used for heating greenhouses from which vegetables and flowers are exported to Europe during the winter months. All systems under development take advantage of the fact that even though Israeli winter nights tend to be cold, there is usually excess heat available for collection during the day. Several researchers call for the collection of this heat in water, which can be recirculated during the evening hours, while others favour its collection in special salt hydrates, which are placed inside double-layered plastic sheets from which the roofs and walls are constructed. The salts change from liquid to solid and back at about 20°C, at the same time releasing large amounts of heat.

Nechemia Meyers

Indian science congress

Pursuit of quality endorsed

Mesra, Bihar

STRINGENT control of higher degrees and special schools for highly gifted children were among the recommendations of the 71st Indian National Science Congress, which met at the Birla Institute of Technology in Mesra last week.

The congress, which resembles a cross between the British Association and a political rally has, since 1976, had a "keynote theme", on which it presents recommendations to the relevant government bodies, and on which it receives an implementation report in the following session. This year's theme was "Quality science in India — ends and means".

This slogan was interpreted in different ways. Professor M.G.K. Menon made an impassioned plea on the serendipity value of basic research. The Prime Minister, Mrs Indira Gandhi, who opened the congress, argued that, since "quality of science" presupposed "quality of life", it implied a joint campaign of scientists against the threat of nuclear war and the subsequent ecological disaster. By the second day, however, the keynote theme had crystallized into how to overcome the prevailing mediocrity of Indian science.

India, as a number of speakers noted, can claim the world's third largest "scientific work-force" in the sense of employed persons with at least a first degree in science. In establishing this human resource, which is vital for the development of Indian agriculture and industry, there has been a concentration on quantity rather than quality of graduates. Furthermore, the low remuneration of young research fellows (around 600 rupees — £40 — a month, or about half the salary of a young graduate entering business or the civil service) means an inevitable wastage on financial grounds. The congress president, Professor R.P. Bambah of Punjab University, maintained that unless research fellows immediately have their stipends doubled, there is no long-term hope for Indian science.

This utopian suggestion did not find its way into the draft "recommendations" which this year seem aimed as much at the scientific community itself as at the government. On the subject of pay, the "recommendations" firmly reject the new scheme of promotion of lecturers to readers and readers to professors on the basis of length of service. This scheme has recently been introduced, under trade union pressure, at a few leading universities, including Delhi, in place of the existing system of promotion by merit as and when a vacancy occurs. The new system, many argued, is a blueprint for apathy and complacency.

One practical suggestion for improving scientific "quality" was that the titles and abstracts of PhD theses should be published together with the names of the

adjudicators awarding the degree. This is intended to counteract the frequently reported slackness and even downright dishonesty in the degree process. At the same time, criteria for higher degrees should be standardized because several universities have acquired a reputation for "easy" PhDs. Government agencies which award research fellowships should establish a joint selection examination, and awards should be made on the basis of merit only — ending the present "positive discrimination" in favour of the disadvantaged. The civil service requirement that certain posts must be filled by graduates should be revoked, thereby reducing the pressure on the universities.

In general, Indian education was felt to

be too examination-oriented — and Professor Bambah's suggestion of "pace-setting" schools for the highly gifted was based on the concept that such young people could better develop their potential in a peer group in which they would learn to work together as a team.

The congress itself had a competitive aspect — competitive awards to young postgraduate scientists selected by the Indian National Academy of Sciences and by the congress. Ironically, the first award announced was for work on RNA. Yet, although it recently hosted a major genetics conference and in spite of the importance of the subject for agriculture, India has been able to establish virtually no university departments of genetics — apparently because of a preponderance of classical botanists and zoologists in the upper echelons of the life sciences.

Vera Rich

Educational computing

Academics gain from competition

Pasadena, California

COMPETITION in the computer industry is nothing new, but computer companies are now competing for an unusual prize — the privilege of giving away millions of dollars of equipment to universities. In recent months the California Institute of Technology (Caltech) has received computers worth some \$1.4 million from Hewlett-Packard and Data General at a 75 per cent discount. Massachusetts Institute of Technology (MIT) has closed a deal with IBM and Digital Equipment Corporation (DEC) under which those companies will donate \$50 million in equipment and support over five years and IBM has announced a programme of grants to 20 universities totalling \$50 million in hardware for the support of manufacturing engineering education. Many other companies are also offering large numbers of small computers to universities at much reduced cost.

The companies' motives are varied. Those offering bulk deals on work-stations may well be hoping to lock in a future market in personal and business computers. The Caltech and MIT deals, however, appear to reflect the companies' need to tap a body of expertise that they themselves lack, particularly in developing educational software and in assembling large integrated computer networks. Charitable donations to universities also yield tax benefits to the companies that substantially reduce their net cost. Educational software will probably be the most tangible benefit to the companies. MIT has agreed to offer IBM and DEC non-exclusive licences for any software developed with their equipment.

The present dearth of serious educational software is universally bemoaned. "The state of the art in educational uses of computers is a disaster", says Steven Lerman, an MIT engineering professor

who directs the programme there. Most educational software is little more than electronic workbooks, which is "the biggest waste of hardware, software, and people's time ever invented". Lerman foresees the use of computers in a variety of truly novel ways that will take advantage of their unique capabilities. "Dry laboratory" experiments could demonstrate chemical titration, thermodynamic principles and the like with the help of sophisticated graphics — and without wasting valuable laboratory time and space.

MIT is itself raising \$20 million for the programme. IBM and DEC, besides donating the machines, are providing software and maintenance support and are lending MIT several full-time personnel.

Here at Caltech, the emphasis is also on educational software development. The university is placing three of the four VAX-class Data General machines it has acquired in the hands of undergraduate students. Eventually, these machines are expected to become the central computers in a network that will include multiple work-stations. Geoffrey Fox, a theoretical physicist who was recently named Caltech's first dean for educational computing, believes it is a mistake, however, to have one computer to each student. Fox says that to get students to use computers and to stimulate the much needed software development requires a diversity of machines with capabilities beyond the "lowest common denominator" that inevitably determines the selection of a personal computer that will be issued to every student. Fox points to one US university that acquired an IBM personal computer for every freshman, only to find that the average computer was used for one hour per day for games and for one hour as a terminal for the university's main computer.

Stephen Budiansky

Greenhouse effect

Energy efficiency may help

Washington

ALTHOUGH a global warming due to a build-up of atmospheric carbon dioxide is inevitable, it can be postponed significantly by improved energy efficiency, according to a new study prepared for the National Science Foundation*. The report, by engineers at Massachusetts Institute of Technology (MIT) and Stanford University, breaks with recent studies by the Environmental Protection Agency (EPA) and the National Academy of Sciences (NAS) in suggesting that practical and effective steps can be taken at once to soften the blow of a warming induced by carbon dioxide. By drawing out to several centuries the time it will take for the atmospheric carbon dioxide concentration to double, society will buy time to develop new energy technologies and to enable agriculture and population distributions to adjust to climatic shifts.

The study found that "realistic energy strategies" based on improvements in energy efficiency could hold the atmos-

The NSF report found that solar and renewable energy technologies are not likely to displace fossil fuels to a sufficient extent to make a significant difference in carbon dioxide emissions in the short run. Nuclear fission, because it will probably be considerably cheaper than coal (especially if acid rain prompts new emission controls on sulphur), is seen as the only energy technology besides improved efficiency with a major short-term effect on atmospheric carbon dioxide.

The study discounted strategies to extract carbon dioxide from fossil fuel combustion and sequester it, for example, in deep ocean waters. This is only one of

several half serious ideas for technical fixes lately advanced. One scheme would add sulphuric acid to the atmosphere — an "artificial volcano" that would cool the Earth, offsetting the carbon dioxide-induced warming — but the price is estimated to be \$50,000 million per year. Another would establish mirrors in Earth orbit to increase sunlight available to high-latitude ocean waters; the increased light would permit the ocean biota to utilize the nutrient excess in these waters, increasing biological uptake of carbon. The capital cost for this venture is estimated very roughly at \$100,000 million.

Stephen Budiansky

*Global Energy Futures and CO₂-Induced Climate Change (by D. Rose, M. Miller & C. Agnew) MIT Energy Laboratory Report No. 83-015, prepared for the Division of Policy Research and Analysis, NSF.

Acid rain

Uncertainty persists in Europe

A REPORT on acid deposition in Britain prepared by the Warren Spring Laboratory for the UK Department of the Environment highlights the need for a long-term programme of measurement on acid precipitation. Asked to examine the distribution in Britain and time trends of acidity in precipitation in Britain and northern Europe, the authors were able to consider only a part of Britain's land surface for lack of measurement stations between which comparisons could be made.

The report therefore calls for a commitment to long-term support for measurement, especially at high altitudes and in urban areas. More use should be made of wet-only collectors (which are not susceptible to contamination by gases and particulate matter) and frequent comparisons between institutions should be made to eliminate differences due to measurement protocols.

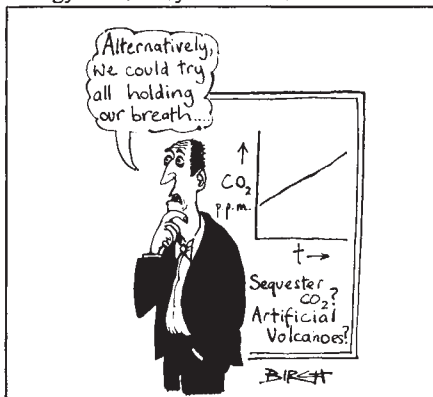
The authors, under the chairmanship of C. F. Barrett of the Warren Spring Laboratory, were unable to reach firm conclusions about recent trends in deposited sulphate, chiefly because of climatic variations from year to year, although nitrate concentrations do appear to have increased since 1957. Looking back further, to the beginning of the century, however, there is evidence that wet depositions of both sulphate and nitrate have increased in approximate agreement with emissions.

From the few sites at which daily analyses are made, the frequency distribution of deposition of acidity seems highly episodic, with typically 30 per cent of annual acidity being deposited in 5 or 6 days. Although acidity cannot be uniquely apportioned between anions, statistical correlations suggest that at most sites studied, sulphuric acid is responsible for about 70 per cent of acidity and nitric acid for 30 per cent. The largest predicted depositions of hydrogen ions occur in Cumbria and parts of Scotland, where levels are of the same order as parts of Scandinavia and North

America. Dry deposition of sulphur, however, is a more local phenomenon related to local sources of sulphur dioxide.

Meanwhile, the European Economic community's Council of Ministers has been unable to reach agreement on a proposal from the Commission of limiting air pollution from industrial plant. The proposal, which is largely procedural, ran aground over an article on emission standards which entailed the eventual possible use of community-wide emission standards, to be introduced by a two-thirds majority. The Commission was prepared to drop the offending article, but the Netherlands objected: Britain and Ireland, in contrast, could not accept the use of such standards in any circumstances. A second much more specific Commission proposal on limiting air pollution from combustion plants, which would have important consequences for power stations, specifies targets for reductions in emissions from both old and new plant, but leaves the options open to governments on how it should be implemented. This proposal has not yet been formally discussed by the council, but is likely to generate an unusual amount of heat.

Tim Beardsley



pheric carbon dioxide concentration to 420 p.p.m. by the year 2050, up from 340 p.p.m. at present. Based on the recent experience of the United States and other industrialized countries, the report concludes that energy use per real dollar of gross national product can be expected to fall by 1 per cent per year.

"It makes sense to develop strategies for reducing fossil fuel carbon emissions, rather than relying solely on research to narrow uncertainties and/or ameliorative measures such as building dikes and developing new strains of 'greenhouse-resistant' crops", the report says. EPA last year suggested that there was little point in trying to halt the carbon dioxide build-up; the NAS report suggested that uncertainties in model predictions (which point to an average global warming of a few degrees centigrade for a doubling of atmospheric carbon dioxide) preclude any immediate steps to halt the build-up.

The NSF study, however, noted that increased energy efficiency, apart from being the only practical route to slowing the build-up of carbon dioxide, is economically attractive in its own right.

Scientific correspondence

Careful readers will have noted that correspondence is now published under two headings — the familiar general correspondence (see pages 103 and 104) and "Scientific Correspondence" (see pages 115 and 116). The objective is to provide a means by which readers can raise points of a rather technical character which are not provoked by articles or letters previously published (where "Matters Arising" remains appropriate).

Contributions to Scientific Correspondence will not usually be formally refereed. □

UK research councils

Astronomers' future undecided

CONTROVERSIAL recommendations are expected from the Manpower and Site Review Panel of the UK Science and Engineering Research Council (SERC), whose proposals have now been sent to SERC's Astronomy Space and Radio (ASR) board for discussion at its meeting later this month. The panel was instructed to find ways of saving more than £1.5 million from the ASR board's budget for the next financial year (1984-85). Whether this has been done by recommending radical changes or by piecemeal cuts, the board will face some painful decisions.

The panel, under the chairmanship of Professor A.P. Willmore, was asked to review the need for each of three establishments — the Rutherford Appleton Laboratory, the Royal Greenwich Observatory (RGO) and the Royal Observatory in Edinburgh (ROE) — taking into account the fact that much of the

United Kingdom's ground-based astronomy will in future be carried out by remote control at facilities in Hawaii (Mauna Kea) in the Canaries (La Palma).

The terms of reference were given considerable sting by the added injunction that the ASR board must save £1.5 million or so. (The board spent about £6.5 million in grants in 1982-83.) RGO had already taken a battering last year in a review of SERC's institutions by a cost-cutting unit of the British Government under Sir Derek Rayner. That report, hotly disputed by RGO staff, recommended that Herstmonceux Castle — a major part of the RGO site — should be sold off (see *Nature* 304, 5; 1983) and that a merger of RGO and ROE should be considered.

When the Willmore panel was set up, it was widely thought by the UK astronomical community that SERC was indeed intending to merge RGO and ROE. The

matter is variously seen as a battle between satellite and ground-based astronomers over the distribution of funds, between SERC's head office and its institutions, or between SERC establishments and universities. Still others fear that the geophysical research supported by the ASR board (mainly concerned with the upper atmosphere and the magnetosphere) will be sacrificed to preserve major astronomical projects such as the millimetre-wave telescope on Mauna Kea and ROSAT, an X-ray satellite supported mainly by West Germany but with UK participation.

Not surprisingly, SERC is keeping its cards close to its chest. Indeed, the proceedings of the Willmore panel have been so secret that its members have not been allowed to remove documents from SERC's offices. The committee's members hold strong and divergent views and it is unlikely that they will make a unanimous radical recommendation. But whatever the panel has decided, the ASR board may choose to ignore it, or instruct the panel to think again.

Philip Campbell

Biotechnology

Gamma interferon race begins

Boston, Massachusetts

GENENTECH has won the race to begin US clinical trials of gamma interferon synthesized in recombinant bacteria. The trials of the drug produced by the California biotechnology company began last month under the direction of a pioneer in clinical interferon research, Dr Jordan Gutterman, a haematologist at the M.D. Anderson Hospital and Tumor Institute in Houston, Texas, one of the largest cancer centres in the United States.

Several biotechnology and pharmaceutical companies have a big stake in the success of gamma interferon, which they hope will be an anticancer drug. The competition is fierce and close: Biogen has been carrying out clinical trials of gamma interferon in Europe and Japan since September 1983. As well as demonstrating gamma interferon's antiviral action, pre-clinical trials by several companies have demonstrated significant antitumour activity, and greater stimulation of the immune system than by other interferons. So gamma interferon could make a good partner for other immunosuppressive chemotherapy drugs.

Gutterman has been treating patients with natural alpha interferon for the past six years, and began trials with cloned alpha interferon three years ago. He and his co-workers reported last week in the *New England Journal of Medicine* (310, 15; 1984) that each of seven patients suffering from hairy cell leukaemia — a slow-growing B cell neoplasm for which there is no consistently effective treatment — were brought into complete or partial remission by injections of natural alpha interferon; Gutterman's group has recently

begun extending these experiments with cloned alpha.

The Texans are looking forward to rapid progress with their tests of gamma interferon; their experience over the past year with the scarce and expensive natural product is expected to help them rapidly to determine the best way to administer the drug — an important point, because intramuscular injections of natural gamma interferon have been found to be cleared rapidly from the body. Gutterman's group has laboriously had to determine the

proper procedure for administering the drug by a prolonged intravenous infusion.

Gutterman is very excited about the possibility that gamma interferon may potentiate the effects of alpha. Despite his success with hairy cell leukaemia, he has found that only a quarter of the patients with melanoma or breast cancer respond to alpha interferon treatment. Substantial *in vitro* and *in vivo* data using mouse and human interferons have shown synergistic antiviral and antitumour effects, and Gutterman hopes that human malignancies may respond similarly. Combined trials may begin in a year or so, once the pharmacology of gamma interferon is worked out.

Christopher Earl

Nature index of biotechnology stocks

12-Month high	12-Month low	Company	Close previous month	Close 30 Dec.	Change
23 1/4	9 3/4	Biogen (Switzerland)	12	10 3/4	-1 1/4
6 1/4	1 5/8	Bio-Logicals (Canada)	2	1 5/8	-3/8
16 1/8	7 1/4	Bio-Response (USA)	11 3/4	10	-1 3/4
19	10 1/2	Cetus (USA)	12 3/8	10 7/8	-1 1/2
15 1/2	6 1/2	Collaborative Research (USA)	8	6 3/4	-1 1/4
39 7/8	15	Damon (USA)	17 3/8	15 3/8	-2 1/4
34 1/4	16 3/8	Enzo-Biochem (USA)	24 1/4	25	+ 3/4
18 7/8	8 5/8	Flow General (USA)	11 1/4	9 1/2	-1 3/4
49 3/4	25 7/8	Genentech (USA)	36 1/2	34 1/2	-2
17 3/4	7 7/8	Genetic Systems (USA)	10 5/8	8 7/8	-1 3/4
23 1/4	12 1/2	Genex (USA)	15	13	-2
31	18 3/4	Hybritech (USA)	22 1/2	18 3/4	-3 3/4
22 1/4	10 7/8	Molecular Genetics (USA)	12 7/8	12	-7/8
23 1/4	10 1/2	Monoclonal Antibodies (USA)	14 3/4	12	-2 3/4
73 1/2	42	Novo Industri A/S (Denmark)	65	56	-9
30 1/4	13 3/8	Pharmacia (Sweden)	21	19 3/8	-1 5/8

Closing prices are for the last Friday of the month. For over-the-counter stocks, bid price is quoted; for stocks on the American and New York exchanges, the transaction price. *Nature's* weighted index of biotechnology stocks stood at 175 on 30 December, compared with 196 a month earlier. Data from E.F. Hutton, Inc.

Chile under Pinochet

SIR — I have read (*Nature* 306, 11–12; 1983) that Chilean universities have not only survived Pinochet's dictatorship but, due to a mixture of despotic benevolence, the existence of a United Nations Regional Program for Post-Graduate Training and the quality of young scientists nurtured by their seniors, Chilean science has in fact been revived. The article exudes an aura of optimism for the future, ending with a reflective statement about the necessity for pragmatic objectives and the dismissal of "over-ambitious utopian dreams and violence".

I read the article with dismay. Certainly, history has taught me that tyrants have at times been generous patrons of the arts and conspicuous builders, but my naive heart would like to believe that tyranny kills all it touches. Furthermore, the article is disturbing because its anecdotal account of hardship falls short of picturing the excesses of Pinochet's murderous tyranny, and because the reported relaxation of the regime is hard to reconcile with reports from Amnesty International of repression, torture and assassination still rampant after 10 years of dictatorship. "Between 1973 and 1979, hundreds . . . disappeared after being arrested by the security forces . . . coordinated by the Direccion de Inteligencia Nacional (DINA). Disappearances and killings . . . remain officially unexplained. Since 1977, a number of alleged members of banned political parties and organizations . . . have died in the custody of the CNI (Central National de Informaciones, Chile's secret police), in circumstances which indicate that they may have died after torture or . . . [been] killed by other means." (*Political Killings by Governments, An Amnesty International Report*, 20–21; 1983).

Let us put the survival issue into perspective. By the time the medical school community was taken to the soccer field, the university had already paid a grievous price — the disappearance of faculty members who were underground, exiled, or already dead. Later on, economic constraints brought the "ousting of many people and early retirements". The survivors were those silenced by fear and various collaborators — those people with deep-rooted rightist convictions and those pragmatic enough to accommodate their principles to the establishment. On the one hand, there was repression and on the other, compliance and consensus. How much consent is hard to evaluate, but guidelines for suppression and removal of academic personnel to be effective after the coup were clearly spelled out in a document signed by a pediatrics professor of the University of Chile.

The subject of the "open and constructive discussion" in Chilean universities which are more open and constructive than those that occur "in other countries

respected for their democratic governments" is not spelled out. I must then conclude that despotism has consented not to be too despotic for people to discuss — allow me to guess — the sodium pump, quantum statistics, the cost of paper clips, promotions and fund allocations. I strongly believe in academic freedom but the liberty that we are told Chilean scientists enjoy occurs simultaneously with hardship upon the Chileans that cannot be silenced nor should it be forgotten. It consists of 50 per cent unemployment in the shanty towns, it consists of the hunger of the poor, desperate citizens fighting with stones against police tear gas and bullets, people still being tortured and in prison; it consists of exiled people, it consists of the dead.

I hope that Chilean transition to democracy will occur soon and that it will be obtained without further violence. However, as long as the aims of a few are placed above the elemental rights of the enslaved many, the germ of violence will be there. The preservation of an ivory tower for a score of people cannot be the constraint to justice and equality. Justice and equality must not be utopian dreams.

ELSA BELLO-REUSS

*Department of Physiology and Biophysics,
Washington University School
of Medicine,
St Louis, Missouri 63110, USA*

Creation or evolution

SIR — In the article "Creation versus evolution: no middle way" (*Nature* 305, 571–574; 1983) George M. Marsden presents an interesting historical review of the roots of the creationism/evolution controversy in the United States, but slights its most important aspect: the science education issue. Because of pressure from antievolutionists, the amount of evolution covered in pre-college science books is documentably declining, or, even worse, is being accompanied by "theories" that the world was produced in six 24-hour days about 10,000 years ago.

There is value in Marsden's reminder that scientific creationism is more than a scientific issue. Proponents are genuinely worried that civilization and the moral order are threatened by modern antisupernaturalism. Because they define creationism narrowly, as Marsden documents, they see evolution as creationism's antithesis. Opposition to evolution thus becomes a moral charge. Textbook censors like the Gablers in Texas discourage discussions of the philosophic implications of various world views because they lead to "relativism", "humanism" and other nasty ideas. They therefore demand that creationism be presented in *science*, not history, social studies or philosophy classes. Unfortunately, there is no scien-

tific evidence for (and much against) ideas such as obligate carnivores being former vegetarians (in order to have lived compatibly on Noah's Ark).

Whether or not evolution is "a key mythological element in a philosophy that functions as a virtual religion" (antisupernaturalism) is beside the point. Indeed, evolution *can* be the basis of a philosophical world view, but conceivably so can heliocentrism, or even photosynthesis. Evolution, heliocentrism and photosynthesis can be presented without philosophical trappings — it is difficult to imagine the latter two being taught otherwise. The facts are that photosynthesis occurs, the Sun is the centre of the Solar System, and change through time (evolution) has occurred. Children need and deserve to be taught this. One can be aware of the role of history and culture in science and still insist that science be taught in the science classroom.

EUGENIE C. SCOTT

*Department of Epidemiology and
International Health,
University of California,
San Francisco, California 94143, USA*

Research Corporation

SIR — Research Corporation takes exception to the article published in *Nature* on 13 October, 1983 (p.569) which alleges mistakes in the handling of a non-isotopic method for detecting gene sequences invented by David Ward of Yale University. That Yale did not "bungle" the technology transfer process is demonstrated by that institution's successful programme of licensing it to industry for development and commercial use.

Research Corporation, a nonprofit foundation that evaluates, patents and licenses inventions for a number of colleges and universities, was erroneously identified as having performed a review of Dr Ward's discovery. In fact, Yale did not submit the invention to Research Corporation, and this organization did not evaluate it or offer any opinion as to its worth.

As a further comment, and in contradiction to writer Budiansky's assertions, Yale University continues to maintain an invention administration agreement with Research Corporation, which now bases the larger part of its Invention Administration Program in Tucson, Arizona (not New York).

JOHN P. SCHAEFER

*Research Corporation,
6840 East Broadway Boulevard,
Tucson, Arizona 85710, USA*

● We apologize for the error. Contrary to what was implied by Yale officials, Research Corporation did not see the original invention. Dr Schaefer is also correct in saying that an arrangement between Yale and Research Corporation continues; it is, however, non-exclusive, and Yale expects largely to assume these functions itself with the establishment of its own office — Editor, *Nature*.

Paying for high-energy physics

SIR — Your article "Paying for high-energy physics" (*Nature* 1 December 1983, p. 421) gives numerous false impressions.

The argument that it was not really necessary to observe the $W^\pm$ and Z^0 bosons, because "the [Salam-Weinberg] theory [of weak interactions] . . . may . . . have been held to have been confirmed by what was learned (at CERN) in the 1970s about neutral currents" is wrong. The neutral current data, in common with all earlier information about weak interactions, are consistent with a force of zero range. It is true that those data are also consistent with the Glashow-Weinberg-Salam (GWS) theory, in which the weak interaction is due to the exchange of the W and Z bosons, thereby having finite range $\hbar/m_Z c \sim 2 \times 10^{-16}$ cm. However, the existence of such bosons was only a speculation until they were discovered at CERN earlier this year. The theoretical arguments for the existence of some structure in weak interactions at short distances were rather compelling but the GWS theory was not unique in fitting the earlier data. There were rival theories of the same type with additional bosons, and also very different theories which predicted W and Z bosons but as composite rather than fundamental objects.

The fact that the masses of the W and Z agree with the GWS predictions, within the errors, is certainly not sufficient to establish their true nature. Indeed, the experiments have reported some tentative evidence for the decay $Z^0 \rightarrow e^+ e^- \gamma$, which cannot be explained by the GWS model. If confirmed, this radiative Z^0 decay might perhaps be accommodated by models in which the W and Z are composite. In any case, these anomalous events may indicate quite new physics to come. The fact that the GWS model is still potentially so vulnerable shows that we are still at the beginning of this chapter of the story.

Next, the 2,500 university physicists who use CERN facilities each year go there because it is their deep wish to be involved. There would be many more if there were funds to support them. It must also be remembered that this subject holds a very deep fascination both for the general public and especially for the young scientist still at school, as becomes very clear from interviewing for university entrance. What is going on at CERN is one of the most exhilarating enterprises mankind has known, a great flight of the creative human spirit to determine in full detail what a strange world we live in, and it is not surprising that so many physicists wish so profoundly to take part in this demanding work. You ask what they would be doing if CERN were not there: the answer is simply — "something less significant and much less interesting".

The insinuation that the advanced techniques used at CERN are just borrowed from American laboratories is simply false.

The proposal to transform the existing CERN Super Proton Synchrotron into the Proton-Antiproton Collider, with which the W and Z were produced, provides the most obvious counter-example. The crucial technique of stochastic cooling of the antiproton bunches — the breakthrough which made this step to much higher energies possible — was proposed and developed at CERN: it has now been taken up by the Americans for their Tevatron collider under construction at Fermilab, near Chicago. CERN also proposed and built the first proton-proton collider (ISR). Furthermore, CERN is notable for its development of new experimental methods; the neutrino horn, multi-wire proportional drift chambers, the development and use of large liquid argon and/or uranium calorimeters, the ring-imaging Cerenkov detector and the ISIS detector (a trend-setter in the problem field of particle identification at high energies) were all pioneered at CERN, to mention just a few outstanding items. Under the pressure of experimental needs, substantial technical developments have been made at CERN which have proved beneficial outside in fields as diverse as electronics and shipbuilding. It would be healthy for this country if more UK engineers were to gain practical experience by taking part in the advanced engineering work going on at CERN all the time.

We agree that the time may be ripe for intercontinental collaboration in the 1990s. For the present, the Americans feel they may be able to build a superconducting collider on their own. However, we cannot anticipate *any* time when experiments will be unnecessary; "cheap theory" alone will not be able to account with certainty for the microscopic structure of the world in which we live, without the aid of experiment.

R. H. DALITZ

C. H. LLEWELLYN SMITH

Department of Theoretical Physics,
University of Oxford,
Oxford OX1 3NP, UK

SIR — The "News and Views" item "Paying for high-energy physics" (*Nature* 1 December 1983, p. 421) contains so many inaccuracies that it is difficult to know how to reply adequately at reasonable length, but perhaps I may attempt to correct some of its more misleading assertions.

Comparing the SF 680 million CERN annual budget with the materials cost of the new electron-positron collider, LEP, is not comparing like with like. The annual budget is divided almost equally between staff costs — and many of the staff are engaged in building LEP at present — and materials costs. The materials budget this year is being spent on running five accelerators (one of which operates in two very different modes), and in building LEP. Electricity consumption is appreciable, but

not outrageous for a large scientific laboratory whose facilities operate round the clock, seven days per week for much of the year. Indeed, it has been estimated recently that stopping all of CERN's accelerators for a complete year would save a total of just SF 25 million in electricity charges — 3.7 per cent of the annual budget.

The super proton synchrotron will not be closed when LEP is completed. It is required as an injector for LEP and, more importantly, is expected to be used in the LEP era by up to half of the European high energy physics community for both fixed target and colliding beam experiments.

It is both scientifically unsound and unnecessarily churlish to attempt to run down CERN's recent triumphs in detecting the W and Z bosons. Maintaining that the observation of neutral currents by CERN a decade ago was a sufficient test of weak interaction theory is as sensible as claiming that the Venus de Milo may be adequately appreciated by studying the shadows cast on the floor by the spotlights that illuminate it. Continuing the analogy, the recent experiments have provided us with a clear silhouette, in remarkable agreement with the descriptions furnished by those distinguished intellectual tourists, Glashow, Salam and Weinberg. We are almost persuaded that they have been telling us the whole truth, but LEP should provide the detailed colour photographs that will settle the issue beyond doubt.

The article poses the question of what CERN's 2,500 visiting scientists would be doing if CERN did not exist. The most plausible answer is: "Working around the world in groups too small to be effective, lacking contact with their peers and the stimulation of direct competition, performing second-rate research with inadequate facilities".

D. C. IMRIE
Chairman,
SERC Sub-Committee on CERN,
Department of Physics and Astronomy,
University College London,
London WC1E 6BT, UK

No thanks — Anon

SIR — Here is a formal "You're welcome" to all those authors who have recently included in the Acknowledgements section of their published papers a line thanking me, an Anonymous Reviewer, for critical comments on their manuscript. I suggest, however, a moratorium on this practice, since thanking me in print is a not particularly informative use of expensive journal space. If authors really feel that they must show their appreciation for an exceptionally helpful anonymous review, they can send me a personal note of thanks via journal editors. Then at least I'll know that it was my comments, and not those of some other Anonymous Reviewer, that were so appreciated. *Name withheld by request*

● *Nature's* policy is not to thank Anonymous Reviewers — when the misdeed is spotted.

Antarctic mining regime at risk

John Maddox

In Washington next week Antarctic Treaty members will wrestle with mineral exploitation. Their task would be easier if territorial claims were subsumed in a wider framework.

LAST week, it was announced from New Delhi that an Indian expedition of 83 people had reached the coast of Antarctica in a chartered Finnish ice-breaker. Next week, in Washington, there will be a special meeting of the Antarctic Treaty powers to resume the drafting of a convention on the exploitation of mineral resources in Antarctica. The two events are strictly independent, but in the minds of delegates gathering in Washington, there will seem to be an important connection. For the accession of India to the Antarctic Treaty last September (together with Brazil) seems to have reminded the parties to the treaty that change is close at hand.

Anxiety about the future of the treaty has been mounting for several years, and certainly since the Law of the Sea Conference reached its abortive conclusion in 1982. The principle that the sea-bed should be regarded as the "common heritage of mankind" was rejected by several governments, that of the United States chief among them, so that the status of the resulting treaty is in doubt. But the earnest advocacy of the doctrine of common heritage by the governments of many developing countries, and its obvious applicability to the Antarctic, has inevitably cast doubt upon the durability of the arrangement under which the Antarctic Treaty vests responsibility for the management of a whole continent in the governments that agreed to give themselves that responsibility at the end of 1959.

In its essentials, the difficulty springs not from the Law of the Sea but from the reasons why the negotiation of that treaty was frustrated—the general unwillingness of newly formed governments to fall in with international arrangements previously agreed. Uppishness of this kind has become a familiar irritant of established governments, in the United Nations and in other international gatherings. But if the Law of the Sea were less like a dead letter, it would present the Antarctic powers with a further and more particular difficulty, for the sea-bed provisions of the Law of the Sea Treaty defined as the common heritage of mankind even those parts of the oceanic sea-bed lying south of 60°S, the region in which the Antarctic Treaty applies.

Mercifully, this potential conflict between the two treaty regimes remains an academic matter. So, superficially, the Antarctic Treaty itself may seem to be. What can be the value of an international treaty to regulate the uses made of a tract of

territory whose practical value is, for the time being, only marginal? That is too detached a view. The benefits of the Antarctic Treaty far transcend its immediate utility. First, the treaty how nations may be persuaded to bury conflicting interests in the pursuit of a more general common interest, for that reason, is a valuable precedent for further international agreements on more urgent matters, arms control for example. Second, the treaty does serve to regulate matters of great importance to one section of the international community, for it has assured access to Antarctica for those wishing to make scientific observations in the region. Third, and more recently, the treaty has provided a framework within which it has been possible to arrange for the orderly exploitation of marine resources. The ambition among the treaty powers to conclude a similar agreement on mineral resources is natural enough, and laudable as well, but more contentious.

To the scientific community, the origins of the treaty are also important because the treaty is a tangible legacy of the International Geophysical Year (IGY), the eighteen-month-long period of co-ordinated geophysical observation from 1 July 1957 to 31 December 1958. Antarctic exploitation was inevitably a conspicuous part of that cross between a planned research programme and a public relations campaign. In the climate of the late 1950s, with the cold war dragging to its end, the notion that the only substantial part of the Earth's solid surface free from weapons of any kind should be permanently demilitarized won ready consent.

In reality, there was little choice. Open access to the Antarctic by scientific expeditions during IGY had been agreed by the states with territorial claims on Antarctica but the declaration by the United States and the Soviet Union that they would continue with Antarctic research after IGY put the claimants in a fix, as did the continuing disputes among those with overlapping claims.

All with an interest in the Antarctic discovered a common interest in postponing a resolution of these difficulties, and a common fear that if they failed to agree among themselves, matters would be taken out of their hands. The result is that the negotiation of the Antarctic Treaty, on the initiative of the United States, was accomplished in just over eighteen months.

The treaty itself is a simple document. It

declares that "in the interests of all mankind", Antarctica shall "continue forever to be used exclusively for peaceful purposes" and that international co-operation and "freedom of scientific investigation in Antarctica" typified by experience during IGY should also continue. Its provisions forbid the establishment of military bases in Antarctica, the testing of weapons (including nuclear weapons) and even the disposal of radioactive waste. Access to all parts of Antarctica is assured, but the twelve original parties to the treaty are free to inspect activities undertaken by other governments—a right exercised almost exclusively by the United States. Nothing is said about the duration of the agreement except that after thirty years from ratification (completed on 23 June 1961), full parties should be free to ask for a conference to review the operation of the treaty and, by implication, to consider its future.

The twelve founder members were a motley crew. They included the two superpowers of the 1950s, Khrushchev's Soviet Union and Eisenhower's United States, a group of European states with traditional interests in Antarctic exploration (Belgium, France, Norway and the United Kingdom), Japan, and a group of states in the Southern Hemisphere including three former British dominions (Australia, New Zealand and South Africa) and Argentina and Chile, neighbours in South America and rivals in Antarctica.

Both the United States and the Soviet Union have eschewed claims of sovereignty in Antarctica. Britain and its former dominions consider, however, that they have legitimate claims on parts of the continent, as do Argentina and Chile. The Norwegian claim (predating Amundsen's trans-Antarctic journey) seems not to be a live issue, but Britain, Argentina and Chile claim sectors of Antarctica which overlap with each of the other two. It is not clear whether France has made a valid claim to sovereignty or has merely reserved the right to do so, but the issue has been put firmly in abeyance by the Antarctic Treaty, under which claimants and non-claimants agree that claims to sovereignty will be neither advanced nor liquidated.

But if the moratorium on the issue of Antarctic sovereignty has not made that contentious issue go away. For one thing, the treaty partly owes its existence to the fear, in the late 1950s, that disputes about

Antarctic sovereignty might get out of hand. Second, the structure of the treaty guarantees continuing precedence for the original signatories.

This is the central issue on which the discontents of other governments have focused. Briefly, the treaty distinguishes between three groups of adherents — the twelve original signatories which are automatically consultative parties, signatories which have qualified as consultative parties by their demonstration of "substantial scientific research activity" in Antarctica and other signatories which have not thus qualified. West Germany and Poland became consultative parties in 1979 and Brazil and India (rather hurriedly) last September. Ordinary signatories of the treaty, which now include the People's Republic of China, number more than a dozen. East Germany, Romania and Uruguay formally accompanied their adherence with a protest at the discriminatory provisions of the treaty. Malaysia, not a signatory of the treaty, has gone further by raising the issue of Antarctica at last year's general assembly of the United Nations.

In the circumstances, the charge that the Antarctic Treaty is a club of self-interested governments cannot easily be resisted. This, after all, is what the treaty implies. The charge that the treaty powers are an exclusive club is, however, unfair. The fact that India was able to qualify last year as a consultative party on the basis of expeditions to Antarctica in two successive summers is a sign that, in the present international climate at least, the consultative parties are not concerned like robins to keep others out. Indeed, the consultative parties have now taken to saying that complainants against the alleged exclusiveness of the treaty have a simple remedy — sign the treaty, mount a couple of expeditions to Antarctica and then become full members. This does not, however, satisfy those who insist that they have no obligation to abide by rules in whose development they played no part.

The broader challenge to the treaty, from states such as Malaysia, based on the doctrine of common heritage and the argument that all members of the United Nations should participate in the administration of Antarctica, will be resisted by many of the consultative parties on more pragmatic grounds. Antarctica, they will insist, has been well managed since 1961, and is likely to be badly managed under a different regime.

On the question of management, the consultative parties are on sure ground. Since 1961, access to Antarctica by scientific expeditions has indeed been open. From the outset, oversight of exploration and of the treaty's requirement that the results of scientific investigations should be freely exchanged has been the responsibility of the Scientific Committee on Antarctic Research (SCAR), a committee established by the International

Council of Scientific Unions which predates the Antarctic Treaty itself. This arrangement, whereby a group of governments has turned over to a non-governmental body such an important scientific advisory function, may well be unique.

SCAR has played an important part in the principal developments within the framework of the treaty since it first entered into force, culminating in the Canberra Convention for the Conservation of Antarctic Marine Resources, which was signed in 1978 and which entered into force in 1980. This had been preceded by two other measures concerned with the conservation of Antarctic flora and fauna (promulgated in 1964) and Antarctic seals, signed in 1972 and ratified six years later.

Like the Antarctic Treaty itself, these measures seek to reconcile conflicting interests by verbal ingenuity. Thus the convention on marine resources does not include the word "exploitation" but "rational use" is implicitly defined as a part of conservation. The result is that the convention on marine resources may be understood as a fisheries convention by those who wish to hunt for krill (chiefly Japan and the Soviet Union). Exploitation is not yet on such a scale as to call in question the restrictions that may in due course be recommended. Meanwhile, SCAR (and its companion committee SCOR, where "O" means oceans) is mounting a detailed investigation of the ecological balance between krill and marine animals in Antarctica.

The exploitation of mineral resources is bound to be more contentious because it touches directly on the issue of sovereignty. The need for an agreement of some kind has been a recurrent theme at consultative meetings since the 1970s. The meeting next week in Washington will be seeking to build on what has been agreed so far — that the need for an understanding is urgent not so much for economic reasons (for nobody suggests that mining anything would be practicable in the near future) but to defend the Antarctic Treaty regime against its critics and that the fragile Antarctic environment must at all costs be safeguarded. Paradoxically, even the non-claimant members of the treaty acknowledge that one of the best defences against the argument that Antarctica should be thrown open to allcomers is that some of them have claims to sovereignty.

The crucial question to be decided is the mechanism by which the exploration and exploitation of Antarctic minerals should be licensed. The consultative parties seem to have agreed that there must be an "external accommodation" in the sense that outsiders should be free to participate provided that they demonstrate practical activity, not merely academic interest, in the discovery of minerals and agree to abide by the rules.

The complementary "internal" accommodation, intended to reconcile the

interests of claimants and non-claimants, is causing more trouble. The intention at present seems to be that there should be some central body with responsibility for mineral exploitation in Antarctica as a whole together with a number of sub-authorities. The hope seems to be that claimant governments would continue to regard involvement with the management of Antarctic resources, perhaps with special responsibility in areas on which their claims have been lodged, as a sufficient recompense for not flying their national flags too ostentatiously.

Legally, there is an important precedent for such an arrangement — the treaty signed in the 1920s under which Spitsbergen (Svalbard), east of Greenland, was acknowledged to be the sovereign territory of Norway, which required that Spitsbergen be demilitarized and which provided for the exploitation of mineral resources (chiefly coal) by other parties. Poland and the Soviet Union pay a modest royalty on what they mine to cover the cost of administration.

Politically, were it not for the overlapping claims of Argentina, Chile and the United Kingdom, the transfer of such a system to Antarctica would be relatively straightforward. Non-claimants could join with the claimants in the detailed administration of sectors of Antarctica, thus satisfying the *amour propre* of all parties and postponing the issue of sovereignty. The snag is that the three overlapping claimants are in no mood to settle on a basis for joint administration of any part of Antarctica — Argentina and Chile cannot agree even about the disputed part of Tierra del Fuego, Argentina and Britain have the Falklands war behind them.

This is why the circumstances call for a new departure. The Antarctic Treaty powers are right to calculate that an external accommodation will not be possible while they cannot agree among themselves on how to regulate their affairs in Antarctica. Against the background of the past three decades of successful co-operation in Antarctica, and with the knowledge that however vast may be the resources hidden beneath the ice-covered continent, the economic value of Antarctic minerals will always be marginal, there is a need for a face-saving form of words.

The ideal would be that the claimants should abandon their claims, but that is too much to ask. A second best is that the claimants should agree that, for some specified period (say a century), their territories should be committed to the trusteeship of others — the twelve original signatories of the Antarctic Treaty for example. That, after all, is the reality of what has been happening since 1961. It would be infuriating if the rewards of two decades of constructive work should now be put in hazard because the claimants, none of which could hope to exercise its sovereignty, cannot swallow such a meagre mouthful of pride. □

From Santorini to Armageddon

North American tree-ring data confirm that volcanic eruptions can cause climatic change. So would nuclear war, but to a degree not yet calculable.

THE general concern with cataclysmic happenings on the surface of the Earth requires no special explanation. In a more or less orderly world, people are naturally disquieted by the notion that their lives may be disrupted (or even ended) by events over which they have no control. For if the collision of the Earth with an asteroid at the end of the Cretaceous may have extinguished the dinosaurs and many other kinds of species, may not less spectacular but more probable events abruptly change the course of history without warning, breaking the comfortable expectation of continuity with the future that we now enjoy?

That, no doubt, is part of the reason why there is such widespread alarm about the risk of nuclear war. The same applies to the more recent general preoccupation with the results of purported recalculations of the long-term consequences of a nuclear war, those for example represented by the summary of a calculation now published by Professor Carl Sagan and a group of associates (*Science* 222, 1283; 1983). They assert that after a nuclear war, the Earth's atmosphere would be so laden with dust that solar insolation of the surface would be drastically reduced, thus also reducing surface temperatures and perhaps even reducing the rate of photosynthesis to below that required for survival.

The issue is not as academic as it might be made to seem by the remark "We'll all be too dead to care". That might be true of the populations of the principal combatants and their neighbours, but populations in the other hemisphere, while damaged by fall-out, would still have a vital interest in the feasibility of survival. So if there are long-term climatic consequences of nuclear war, it is important to know what they are. Those wishing to follow the argument might prudently begin not with the article by Sagan *et al.* but with that of Lamarche and Hirschboeck from the University of Arizona in this issue of *Nature* (page 121).

That rich tale, fascinating in its own right, also has a bearing on the question raised by Sagan *et al.* What Lamarche and Hirschboeck have done is to correlate with the known occurrences of major volcanoes the occurrences of frost damage rings in bristlecone pines, the long-lived species by means of which dendrochronology in eastern California has been extended back for nearly 5,500 years. The results are striking. Frost-rings in bristlecone pines do indeed correlate well in time with the occur-

rence of major volcanoes, among which Krakatoa (1883) is, so to speak, the type-specimen.

Within the past century, during which historical records approach in accuracy the tree-ring calendar, the data show that frost-rings follow eruptions more frequently than if the two kinds of events were independent. But during this period, there were summer cold snaps in the western United States not attributable to eruptions (1941 and 1886 for example) and, in the earlier period, many major eruptions that appear not to have caused frost damage.

That is what would be expected. The veil of dust from a volcanic eruption is only one of many causes of climatic variation, while eruptions differ markedly in character, some yielding only short-lived dust clouds.

Several features of the study now described are nevertheless especially pleasing. First, the notion that major eruptions can cause detectable climatic events is modestly confirmed. Second, the work further supports the view that one of the meteorological consequences of volcanic dust in the atmosphere is (in the Northern Hemisphere) an expansion of the circumpolar vortex in the mid-troposphere and a tendency for the pattern of the stratospheric jet-stream to consist of a more exaggerated sequence of lunges towards low latitudes. That should enhance the confidence of model-builders and support those who have been suggesting that the cold spring and autumn in the Rocky Mountains last year (1983) may indeed have been a consequence of the El Chicón eruption in 1982. Finally, a valuable bonus, there is an objective date (1626 BC) for the volcanic explosion of virtually all of the island of Santorini (then called Thera), one of the traumatic events of late Minoan history.

The bearing of this careful study on the question raised by Sagan *et al.* is only indirect. It does, however, qualitatively confirm one of the chief predictions of the consequences of nuclear war — that dust in the atmosphere can have long-term climatic consequences. But many of the events suggested by the tree-ring study seem to have been exceedingly short-lived. Thus the frost-rings found in California, Utah and Nevada in 1884, the year after the eruption at Krakatoa, seem to correspond with a southerly incursion of cold air from Canada on 9 and 10 September (after a summer that had admittedly been unusually cool).

By contrast, Sagan *et al.* predict much lowered temperatures — on some pictures of the course of a nuclear war, perhaps 20 per cent below normal for 3–6 months afterwards. The chief reason for this huge discrepancy between the consequences of a nuclear war and a volcanic eruption stems from the assumptions made about the quantities of dust carried into the stratosphere by the different events. But everybody, Sagan *et al.* included, agrees that this field is shot through with uncertainty.

The assumptions made about nuclear dust appear, however, to be chiefly responsible for the differences between the new and earlier calculations, that of the 1975 report from the National Academy of Sciences for example. In what is called the "baseline scenario" of a nuclear war (10,000 nuclear explosions with a total yield of 5,000 megatonnes), it is supposed that 1,000 million tonnes of dust would be created, that 80 per cent of that would find its way into the stratosphere and that 8.6 per cent of that would consist of particles smaller than 1 micrometre radius. These assumptions would load the atmosphere with more fine dust than previous calculations have assumed. The new assumptions (and the use made of the one-dimensional model for predicting the attenuation of the obscuring effects of stratospheric dust) are, unfortunately, to be discussed only in a paper still "in preparation". Until that appears, those familiar with the well documented difficulties of measuring the composition and the size distribution of dust even from well studied volcanoes such as Mount St Helens (see for example Brazier, S. *et al. Nature* 301, 115; 1983) will be wondering where the apparently much more precise data about nuclear dust has come from.

None of this implies that such calculations should wait on accurate parameters for all possible variables. It would probably also be common ground between the authors of the latest calculation and their more quizzical readers that the use of one-dimensional models is likely to yield upper limits for the effects calculated. What might, however, reasonably be asked of the authors is that their conclusions, derived as they are from necessarily crude models and uncertain data, should be plainly stamped with the label QUALITATIVE for fear that their apparent precision may prove spurious.

John Maddox

Genetic engineering

Prospects for transforming monocot crop plants

from Richard Flavell and Raymond Mathias

THE past year has at long last seen the successful genetic modification of plants by the insertion of DNA from the test tube — although as yet only tobacco and a few closely related species have been transformed with new active genes¹⁻⁴. Success has come by exploiting the components of the natural DNA transfer system of the soil bacterium *Agrobacterium tumefaciens*, a pathogen of many dicotyledonous species (dicots). During the infection process, the bacterium transfers a part of one of its plasmids, the Ti plasmid, into the host plant chromosomes⁵. What the genetic engineers have done is to place new genes into an appropriate segment of this plasmid, using a combination of *in vitro* recombinant DNA techniques and *in vivo* bacterial genetics. When the bacterium carrying the engineered plasmid is used to infect plant cells, the plasmid segment carrying the new genes is transferred into the plant chromosomes.

The excitement that has rightly accompanied these successes has been tempered by the knowledge that the natural host range of *A. tumefaciens* is restricted only to dicots. So an efficient technique for DNA transfer cannot be used to transform monocotyledonous species (monocots). Yet the genetic transformation of monocots remains a most important goal in plant genetic engineering, for many of the world's major food crops are monocots — rice, barley, maize, wheat and sorghum, for example. The many claims that 'transformed' monocot plants have been obtained after the microinjection or uptake of DNA into immature or mature seeds⁶⁻⁸ have been based on the observation of altered plant phenotypes, but have not been supported by the more definite proofs that can be gained using molecular biology. Much effort is now being directed towards giving such proof.

The successful transformation of tobacco involves the incorporation of DNA into single cells, the selection and culture of transformed cell lines and the regeneration of whole plants therefrom. Few plant species can, however, be taken from single cells through the tissue culture steps as readily as can tobacco and the important monocot crop species have proved particularly recalcitrant in this respect⁹. But while a reproducible single-cell system capable of efficient regeneration into whole plants may still be some way away for the important monocot species, recent developments suggest that excessive pessimism is unwarranted.

In maize, for example, certain cell culture lines produce protoplasts which

readily regenerate to the callus stage^{10,11}, while other lines can be maintained as cell suspension cultures in which aggregates of 20–50 cells can be regenerated to callus and plants with high efficiency¹². Protoplasts from embryogenic cell suspension cultures of pearl millet have been reported to undergo sustained divisions that give rise to calli, embryoids and whole plants¹³. Also protoplasts of barley, rice, sorghum and diploid wheat have been proliferated to the callus stage but without plant formation¹⁴. These developments do not necessarily imply that it will soon be possible to insert DNA into single monocot cells and to regenerate transformed plants efficiently from them, but it may be now worthwhile attempting the microinjection of DNA into the separate cells of a cell cluster capable of regenerating into a transformed plant.

As the regeneration of whole plants from single cells is likely to remain a severe problem for many monocots, other routes for plant transformation are now being explored. These approaches seek to avoid the problems of cell culture altogether by making the gametes the recipients of introduced DNA. The pollen of many species can be germinated *in vitro*¹⁵ and the sperm nucleus can be made a target for the microinjection of DNA as it passes down the pollen tube. If the injected genes are maintained in the nucleus after transfer of the germinated pollen to an ovary, they may be incorporated into the zygote nucleus by the natural fertilization process, resulting in a transformed embryo and potentially a transformed plant. In those species in which the pollen remains viable for only a short period or in which pollen germination *in vitro* is not yet possible, microinjection directly into the egg cell is being explored as an alternative. The principal difficulty in this approach is that the egg is buried in the maternal ovary tissue and is difficult to locate but may be less serious in those few species whose fertilized ovules can be cultured and microinjected after excision.

Little is as yet known, in plants, of the influence of the structure of the potentially transforming DNA which, in principle, can influence the survival of the DNA, the efficiency with which it is integrated into a host replicon and whether it behaves as a replicon. It may be that any piece of DNA introduced into a nucleus has a high probability of being incorporated into a chromosome, as found in mammalian cells¹⁶. The dramatic progress in transforming *Drosophila* by the microinjection of P-elements into suitable recipient genotypes is a striking illustration of the value of

using vectors with special properties¹⁷⁻¹⁹. Complete P-elements encode genes which facilitate their autonomous integration into the chromosomes of *Drosophila* lines not carrying specific repressors of P-element activity. Novel genes spliced into these elements disrupt the essential 'mobility' gene functions, but when such 'inactive' engineered elements are microinjected into *Drosophila* embryos along with unmodified P-elements, the 'trans' acting gene products of the unmodified elements facilitate the chromosomal integration of the engineered element with their new genes.

Transposable elements were first recognized in maize²⁰, and new examples have now been isolated from maize chromosomes by DNA cloning techniques²¹⁻²³ (see, for example, this issue, page 127). These elements are therefore for use in experiments analogous to those which have been so successful for the incorporation of genes into *Drosophila*. The segment of plasmid DNA that is transferred to dicotyledonous chromosomes by *Agrobacterium* is a type of transposable element. Many laboratories are therefore beginning experiments to see if its transposition can be induced in monocot cells when the Ti plasmid is introduced into cells by means other than natural infection by *Agrobacterium*.

The recognition of transformed cells, tissues or plants, is of critical importance in transformation studies, for successful events may occur only at low frequencies. Consideration must therefore be given to the availability of selectable marker genes for incorporation into vectors. Suitably engineered antibiotic resistance genes from prokaryotes have already been used successfully to select transformants from the *Agrobacterium* transformation system and these genes¹⁻⁴, with their proven efficiency, will no doubt be incorporated into monocot vectors. An interesting selection scheme has been developed for transforming cell lines lacking functional alcohol dehydrogenase (ADH) genes: the vector genes will carry a functional ADH gene²³, putative transformants will be cultured under anaerobic conditions where selection operates in favour of cells containing ADH²⁴ and the presence of functional ADH detected easily by a simple staining reaction²⁵.

Taken together, these developments have clearly demonstrated the feasibility of plant transformation in dicots. The progress in monocot cell culture and the expanding interest in the development of alternative methods of DNA delivery, coupled with the availability of Ti-plasmid, monocot transposable elements and selectable markers, suggest that the transformation of at least some monocot species cannot be far off. The need to combine the

Richard Flavell and Raymond Mathias are in the Plant Breeding Institute, Maris Lane, Trumpington, Cambridge CB2 2LQ.

right techniques with the right recipient cells and a suitable DNA delivery system may still keep us waiting a little while longer. We must remember too, that for agriculture, the ability efficiently to insert genes into one monocot crop species is not enough. We must learn how to manipulate all the major crop species. □

1. Herrera-Estrella L. *et al* *EMBO J.* **2**, 987 (1983).
2. Herrera-Estrella, L., Depider, A., Van Montague, M. & Schell, J. *Nature* **303** 209 (1983).
3. Fraley, R.T. *et al* *Proc. natn. Acad. Sci. U.S.A.* **80**, 4803 (1983).
4. Bevan, M., Flavell, R.B. & Chilton, M.D. *Nature* **304**, 184 (1983).
5. Bevan, M.W. & Chilton, M.D. *A. Rev. Genet.* **16**, 357 (1982).
6. Ledoux, L. & Huart, R. *J. molec. Biol.* **43**, 243 (1969).
7. Soyfer, V.N. *et al* *Mutat. Res.* **36**, 303 (1976).
8. Korohoda, J. & Strzalka, K.Z. *Pflanzenphysiol.* **94**, 95 (1979).
9. Potrykus, I. *Advances in Protoplast Research* (eds

- Ferenxzy, L., Farkas, G.L. & Lazar, G.) (Akademiar Kiado, Budapest, 1980).
10. Potrykus, I., Harms, C.T. & Lorz, H. *Theor. appl. Genet.* **54**, 209 (1979).
11. Chourey, P.S., Zurawski, D.B. *Theor. appl. Genet.* **59**, 341 (1981).
12. Green, C.E., Armstrong, C.L. & Anderson, P.C. *Miami Winter Symp.* **20**, (1983).
13. Vasil, V. & Vasil, I.K. *Theor. appl. Genet.* **56**, 97 (1980).
14. Vasil, I.K. in *Plant Improvement and Somatic Cell Genetics* (eds Vasil, I.K., Scowcroft, W.R. & Frey, W.J.) 179-204 (Academic Press, 1982).
15. Raman, K., Walden, D.B. & Greyson, R.I. *J. Heredity* **71**, 311 (1980).
16. Wigler, M. *et al* *Cell* **16**, 777 (1979).
17. Spradling, A.C. & Rubin, G.M. *Science* **218**, 341 (1982).
18. Rubin, G.M. & Spradling, A.C. *Science* **218**, 348 (1982).
19. Spradling, A.C. & Rubin, G.M. *Cell* **34**, 47 (1983).
20. McClintock, B. *Cold Spring Harb. Symp. quant. Biol.* **16**, 13 (1951).
21. Federoff, N., Wessler, S. & Shure, M. *Cell* (in the press).
22. Döring, H.P., Tillmann, E. & Starlinger, P. *Nature* **307**, 127-130 (1983).
23. Peacock, W.J. *et al* *Miami Winter Symp.* **20**, (1983).
24. Shimamoto, K. & King, P.J. *Molec. gen. Genet.* **191**, 271 (1983).
25. Freeling, M. *Genetics* **83**, 70 (1976).

Genetics

The origins of men with two X chromosomes

from Paul Burgoyne

APPROXIMATELY one in 20,000 men have a female (XX) sex-chromosome constitution. The occurrence of XX males is of great interest because it raises the question of whether testes can develop in the absence of the testis-determining locus located on the Y chromosome. Recent work has done much to support the view that XX males have a testis-determining fragment of the Y chromosome attached to one X chromosome, but a group of three related XX men do not fit this explanation. As is often the case, these exceptions are potentially the most informative.

In 1966 Ferguson-Smith¹ suggested that the paternally derived X of XX males may harbour a testis-determining fragment of the Y. He envisaged that the fragment might have been transferred to the X by 'accidental crossing-over' between the X and the Y in the father and referred to the event as an X-Y interchange. Originally the idea of X-Y interchange was founded on the observation that some XX men fail to inherit their father's allele for the X-linked red cell antigen Xg^a. The interchange model explains this by assuming that the Xg locus, which is located at the tip of the short arm of the X chromosome, is lost from the paternal X when it acquires part of the Y short arm containing the testis-determining locus.

In 1979 another dimension was added to the X-Y interchange model when Evans *et al.*² showed a difference in length between the short arms of the two X chromosomes of some but (as others have emphasized³) not all XX men. This was taken as evidence for the addition of Y-chromosomal material to the short arm of the paternal X, and in terms of the interchange model this implied that the accidental crossing-over could be unequal with more Y-chromosomal material being gained than

X-chromosomal material lost. Polani⁴ considers that the organization of the X and Y pairing segments may predispose them to unequal exchanges.

Two recent reports provide further genetic evidence of X-Y interchange. The first⁵ indicates that the steroid sulphatase locus (STS), which, like Xg, is near the tip of the X short arm and escapes X-inactivation, can likewise be missing from the paternal X in some XX males. While most XX men have high levels of STS activity befitting their XX status, in the case described not only was STS activity low but one of the Xs failed to express STS when incorporated in a human-mouse somatic cell hybrid line. The second report⁶, published in this issue of *Nature*, describes an XX male who failed to inherit his father's Xg allele but did inherit his father's Y-linked 12E7 ('Yg')⁷ allele. This is the first genetic demonstration of both the loss and gain of material by the paternally derived X of an XX male.

In view of the weight of evidence for X-Y interchange in the aetiology of XX maleness, it was reassuring to hear a report by Bishop, Fellous and Weisenbach at a recent meeting in Oxford that by using Y-specific probes⁸ they have identified Y-specific sequences in the DNA from all three XX males they have examined (and see this issue of *Nature*)⁹.

The case for X-Y interchange thus seems to be proven, but is this the whole story? Albert de la Chapelle³ has long argued for a heterogeneous aetiology for XX males, and when the extremely rare familial cases are considered there is certainly food for thought. Let us consider two pedigrees; that described by Kasdan and colleagues¹⁰ where there is an autosomal dominant pattern of inheritance, and that of de la Chapelle³ in which transmission through

females formally rules out dominance. The autosomal dominant pedigree has a direct counterpart in the XX sex-reversal mutation *Sxr* in the mouse. Although inherited in autosomal fashion *Sxr* has turned out to be a fragment of the Y which is regularly transferred to the X during meiosis in carrier males^{11, 12}. This intriguing finding was discussed in an earlier article¹³. The important point in the present context is that an apparently autosomal pattern of inheritance can be produced if crossing-over is obligatory in the pairing segment of the X and Y¹⁴. Thus the Kasdan pedigree may, like *Sxr*, be explained by X-Y interchange.

The de la Chapelle pedigree cannot be dismissed so easily. First, the recessive inheritance in my view argues strongly against the involvement of the Y-borne testis-determining locus. More important, in one of the three XX males in this pedigree both X chromosomes appear to be maternal, since he does not express the paternal alleles for Xg (located at the tip of the X short arm) or X_m (located near the end of the X long arm). I believe this unlikely anomaly of X inheritance suggests that this XX male has received from his mother two X chromosomes bearing a recessive 'testis-determining' mutant. Wolf¹⁵ has put forward a model for testis determination in which testis differentiation is prevented in XX individuals by an X-borne repressor locus. When a Y chromosome is present this repressor locus is in some way neutralized so allowing testes to develop. This model allows for the development of testes in XX individuals if both copies of the X-borne repressor locus are inoperative through mutation or deletion. Such an X-borne recessive effect is consistent with the de la Chapelle pedigree, provided that the X-borne repressor locus is inherited in a pseudoautosomal fashion (like *Sxr*). I have previously suggested pseudoautosomal inheritance of this repressor locus in order to explain recessive XX sex reversal in goats¹⁴.

The XX males from these two pedigrees are a must for investigation using Y-specific probes. The XX males from the Kasdan pedigree may well reveal Y-chromosomal sequences, but I for one shall be very surprised if this is also true of de la Chapelle's XX males. □

Paul Burgoyne is at the MRC Mammalian Development Unit, 4 Stephenson Way, London NW1 2HE.

1. Ferguson-Smith, M.A. *Lancet* 27 August, 475 (1966).
2. Evans, H.J. *et al* *Hum. Genet.* **49**, 11 (1979).
3. de la Chapelle, A. *Hum. Genet.* **58**, 105 (1981).
4. Polani, P.E. *Hum. Genet.* **60**, 207 (1982).
5. Wieacker, P. *et al* *Cytogenet. Cell Genet.* **35**, 71 (1983).
6. de la Chapelle, A. *et al* *Nature* **307**, 170-171 (1983).
7. Goodfellow, P.N. & Tippet, P. *Nature* **289**, 404 (1981).
8. Bishop, C.E. *et al* *Nature* **303**, 831 (1983).
9. Guettaen, G. *et al* *Nature* **307**, 172-173 (1983).
10. Kasdan, R. *et al* *New Engl. J. Med.* **288**, 539 (1973).
11. Singh, L. & Jones, K. W. *Cell* **28**, 205 (1982).
12. Evans, E. P. Burtenshaw, M. & Cattaniach, B.M. *Nature* **300**, 443 (1982).
13. Simpson, E. *Nature* **300**, 404 (1982).
14. Burgoyne, P.S. *Hum. Genet.* **61**, 85 (1982).
15. Wolf, U. in *Development and Function of Reproductive Organs*, 260 (Excerpta Medica, Amsterdam).

Applied mathematics

Scattering by fractal objects

from Eric Jakeman

"A FRACTAL is a mathematical set or object whose form is extremely irregular and/or fragmented at all scales." So runs Mandelbrot's definition of the term he coined in an essay first published¹ in French in 1975. He points out that classical mathematics has on the whole been concerned with Euclidean geometry based on the notion of smooth shapes and with the continuous evolving dynamics of Newton, whereas most natural phenomena, objects and patterns are abruptly changing, with structure on many scales. Thus the mathematical description of a whole range of familiar objects ranging from coastlines to trees presents great difficulties. In his essay, Mandelbrot shows how these can be resolved by the concepts of fractal geometry².

To reduce the totality of fragmented objects to a class requiring minimal parameterization, it is necessary to introduce the notion of 'scaling', which refers to the similarity in appearance of an object when it is subject to arbitrary magnification. Regular scaling fractals can be generated by following simple iterative construction procedures. The best known object of this kind is the Von-Koch 'snowflake' curve³ obtained by repeatedly replacing the middle third of each element of the construct by two vertices of the appropriate equilateral triangle (Fig. 1). This leads to an object of finite area but infinite perimeter. Its bounding curve is evidently continuous but does not have a locally well-defined direction (slope or derivative).

Although a number of these strange but regular geometrical constructs are known, the vast majority of natural objects are fragmented in an irregular way and it is in the description of these randomly varying structures that the concepts of fractal geometry are proving most valuable. An early investigation of one such structure is attributed to L.F. Richardson, who found that the length of land frontiers and coastlines appears to increase as the resolution of the measurement improves². Thus, if the length of the coastline of Britain is determined using a 100-km measuring stick, the answer will be smaller than that obtained using a 10-km stick, which will

resolve and therefore have to negotiate the longer distance round smaller scale structures such as river estuaries and inlets.

This observation leads quite naturally to the principle parameter characterizing a scaling fractal curve, for it is found that if the length of the measuring stick is l then the apparent length of such a curve is proportional to l^{1-D} , where D is a number greater than unity, but not necessarily integral, which uniquely characterizes the scaling behaviour of the curve. The value of D is in fact a measure of the density with which the curve fills the space (surface, volume) in which it is embedded and can be interpreted as an anomalous or fractal dimension.

Although Mandelbrot's ideas have diffused rapidly, the most striking applications of his work remain the numerical simulations of random fractal objects such as islands and landscapes which, in spite of being entirely artificially contrived, acquire an uncannily familiar appearance when supplemented by suitable shadowing effects². The ability to recognize the familiar in this context is dependent, of course, on information being carried to the eye by light scattered from the object and there is indeed increasing interest in the properties not only of light but of other frequencies of the electromagnetic spectrum as well as of sound waves which have been scattered by fractal objects.

In fact there is a long history of interest in one multi-scale scattering system — the turbulent medium. Although not referred to as such, a fractal description involving the Kolmogorov spectrum has commonly been employed since the pioneering work of Tatarski in the early 1960s⁴. The full implications of using models of this type have only become clear recently, however, through the quantitative investigation of scattering by simple models of fractal surfaces⁵. Since the slope of a fractal surface is ill-defined, ray or geometrical optics do not apply and the scattering is of an entirely diffractive nature. This results in relatively low contrast intensity (brightness) patterns⁶ that are very different from those generated by smoothly varying surfaces whose lens-like behaviour

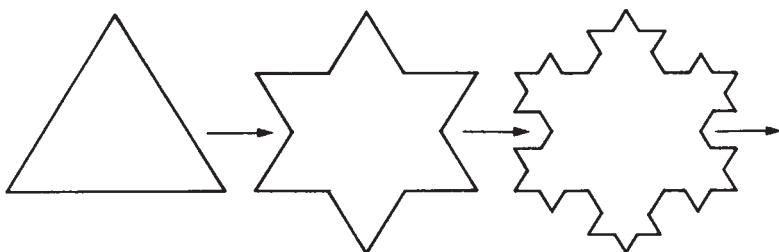
produces bright geometrical features of the kind which can often be seen on the floor of a swimming pool.

In the case of pulsed scattering, the fractal dimension of the surface is represented in the temporal decay of the tail of the return pulse, so that this kind of surface should in principle be easily recognizable and capable of characterization by remote sensing techniques⁷. In practice variations of reflectivity may confuse such measurements, while real surfaces will also have smallest (inner) and largest (outer) scale sizes with a hierarchical or fractal structure in between. To investigate the latter regime, the wavelength of the probing radiation has to be much greater than the height fluctuations over inner-scale sized elements of the surface (so that their scattering effect can be neglected by comparison with that of the larger scales) and the receiving optics must be designed to collect radiation from an area of the surface which is smaller than the outer scale size.

Although many solid surfaces of both microscopic and geophysical proportions appear to be fractal⁸, at least over some limited range of scale sizes, there is evidence that certain fluid systems are better described as objects with fractal slope. The sea surface is one such system, for example. Ray or geometrical optics effects do play an important role in determining the characteristics of the scattered intensity pattern in this case, leading to strong variations in brightness. Since the local surface curvature remains ill-defined, however, lens-like behaviour does not occur and a new kind of geometrical optics or short-wave limit is found which is free from the catastrophes (caustics, focusing points) associated with smoothly varying surfaces. The statistical properties of the scattered intensity remain finite in this limit and can again be related to the fractal dimension⁹.

The full implications of current work on fractal scattering cannot yet be properly assessed but will surely extend beyond the immediate practical objectives of the investigations. Certainly many new and sometimes surprising results are emerging by recognizing the role of fractals in well-established areas of science as well as in the more exotic frontier areas such as dynamical systems and chaos. □

Eric Jakeman is in the Royal Signals and Radar Establishment, St Andrews Road, Great Malvern, WR14 3PS.



The Von-Koch snowflake curve

1. Mandelbrot, B.B., *Les objets fractals: forme, hasard et dimension* (Flammarion, Paris, 1975).
2. Mandelbrot, B.B., *The fractal geometry of nature* (Freeman, San Francisco 1982).
3. Cundy, H.M. & Rollet, A.P., *Mathematical models* (Clarendon, Oxford 1954).
4. Tatarski, V.I., *Wave propagation in a turbulent medium* (McGraw-Hill, New York, 1961).
5. Berry, M.V., *J. Phys. A* 12 781 (1979).
6. Jordan D.L., Hollins, R.C. & Jakeman, E. *Appl. Phys.* B31 179 (1983).
7. Berry, M.V. & Blackwell, T.M., *J. Phys. A* 14, 3101 (1981).
8. Sayles, R.S. & Thomas, T.R. *Nature* 271 431 (1978).
9. Jakeman, E. *J. Opt. Soc. Am.* 72, 1034 (1982).

Galaxy structure

Two kinds of halo gas identified

from Joel Bregman

ALTHOUGH most of the cool gas in our galaxy lies in the galactic disk, it is not entirely confined there and the study of gas above and below the disk (halo gas) has been an especially vigorous area of research during the past few years. Unfortunately, it has frequently been difficult to discriminate between disk and halo gas, for it is not possible to determine distances to gas clouds and we can only look at the halo gas through the disk. In a recent paper C. E. Albert has used a new approach¹, a study of low ionization gas within 8,000 light years of the disk (the lower galactic halo). She has conclusively shown that a considerable amount of neutral (cool) gas exists in the lower galactic halo and that there are two kinematically distinct types of halo gas.

Gas is believed to exist in the halo in several different forms². Weak radio continuum emission has been attributed to synchrotron emission from cosmic ray electrons that may extend up to 30,000 light years from the disk^{3,4}. Plasma at temperatures of millions of degrees may also permeate the halo and be responsible for part of the diffuse X-ray background (0.1–2 keV)⁵. Virtually nothing is known about the vertical distribution of this gas except that it is unlikely to extend beyond 30,000 light years from the disk. Radio telescope surveys have revealed the presence of neutral hydrogen gas at a few hundred degrees kelvin but with unusual velocities^{6,7}; this gas is presumed to lie in the halo as well. Finally, absorption line spectroscopy in the optical and ultraviolet bands has been used to demonstrate the presence of halo gas with a temperature in the range 10^3 – 10^5 K^{8–10}.

Optical and ultraviolet spectroscopy have been powerful tools for analysing the properties of halo gas. By detecting absorption lines caused by gas intervening between the Sun and a background source (a star or quasar), one is able to estimate the column density, vertical distribution, chemical abundance and temperature of the gas. Strong absorption is commonly detected for the ultraviolet resonance lines of C IV and Si IV, which is interpreted as evidence for warm gas (10^4 – 10^5 K) lying between 1,500 and 10,000 light years above the disk^{9–11}. For neutral gas ($< 10^3$ K), a more precise study of the vertical distribution has been made by comparing the total column density (from the 21 cm H I emission line) with the column density between the Sun and a halo star (using the Ly α line of neutral hydrogen)¹².

There is a vast amount of gas above the thin disk of the galaxy (half-thickness about 300 light-years) and about one-third of the total column density of neutral gas is

beyond a height of 3,000 light years. However, little neutral gas exists further than 6,000 light years from the disk. Albert has added to our understanding by comparing the difference in absorption for pairs of stars, one of which is just above the galactic disk and the other in the halo. She shows that the halo is rich in gas and distinguishes between two types of neutral halo gas: a layer of gas similar to disk gas that has a scale height of about 3,000 light years, and higher velocity gas richer in certain heavy elements (Ti, Ca, and Na) that is found only above the disk. Whether or not there is a true difference between the distribution of warm and neutral gas and how these components are related to the million degree gas and the cosmic rays remains to be seen.

The absorption-line gas is rotating, but not precisely with the galactic disk. This has been clearly shown by Albert and others¹³ who demonstrated that the velocities at which absorption against halo

stars is observed are considerably different from these expected for corotating gas. Evidently, these gas clouds have large peculiar velocities with respect to a corotating halo. The data are qualitatively consistent with features of a corona that rotates more slowly than the disk and of galactic fountain models. More data are needed before the various models can be tested critically and more studies along the lines of Albert's would be valuable. □

Joel Bregman is in the department of physics at New York University.

1. Albert, C.E. *Astrophys. J.* **272**, 509 (1983).
2. York, D. G., *Ann. Rev. Astr. Ap.* **20**, 221 (1982).
3. Webster, A. *Mon. Not. R. astr. Soc.* **185**, 507 (1978).
4. Higdon, J.C. *Astrophys. J.* **232**, 113 (1979).
5. Nousek, J. A., Fried, P. M., Sanders, W. T. Kraushaar, W.L. *Astrophys. J.* **258**, 83 (1982).
6. Giovanelli, R. *Astr. J.* **85**, 1155 (1980).
7. Schwarz, U. J. & Oort, J.H. *Astrophys. J.* **101**, 305 (1981).
8. Songaila, A. & York, D.G. *Astrophys. J.* **242**, 976 (1980).
9. Savage, B.D., & de Boer, K.S. *Astrophys. J.* **243**, 460 (1981).
10. Pettini, M. & West, K. A. *Astrophys. J.* **260**, 561 (1982).
11. York, D. G., Blades, J. C., Cowie, L. L., Morton, D.C., Songaila, A. & Wu, C.C. *Astrophys. J.* **255**, 467 (1982).
12. Hobbs, L.M., Morgan, W.W., Albert, C.E. & Lockman, F.J., *Astrophys. J.* **263**, 690 (1982).
13. deBoer, K.S. & Savage, B.D. *Astrophys. J.* **265**, 210 (1983).

Palaeontology

Small companions for early dinosaurs

from Michael Benton

FOSSILIZED cave systems around Bristol, England, have yielded an important fauna of small reptiles dating from the times of the early dinosaurs, about 200 million years ago. Until recently, these were thought to be specialized upland faunas because the animals were rather different from those of the typical lowland dinosaur beds. Strong evidence has now, however, been presented against this view, and new information on the dating and palaeoecology of these remarkable faunas has also been given^{1–6}.

Fossil mammal teeth were first found in the Bristol fissures by Charles Moore over 100 years ago⁷, and the first reptile remains were described⁸ in 1939. These consist of the bones of a small lizard-like animal (*Clevosaurus*), very like the unique living tuatara, *Sphenodon*, from New Zealand. Many more important mammal and reptile fossils have been collected since then from dozens of fossil caves in South Wales and around Bristol. The reptiles include a remarkable small gliding animal called *Kuehneosaurus*, isolated bones of the carnivorous crocodile-shaped *Rileya* and the dinosaur *Thecodontosaurus*, a crocodile, and various as yet undescribed sphenodontids and small archosaurs (crocodiles/dinosaurs/theodontians). The

latest new forms to be described have been *Gephyrosaurus*^{5,6}, a 40-cm long lizard-like animal, and *Planocephalosaurus*, a small sphenodontid³.

The dating of the fissures has always been a problem. Some may be late Triassic (Norian) in age, others Rhaetic, and others early Jurassic (between about 190 and 225 million years ago). Some of the South Wales fissures have been dated as Rhaetic or early Jurassic on the basis of plants, spores and gastropods⁹. One of the fissures near Bristol has yielded more abundant spores which suggest a Rhaetic age¹. The main fissure sites, and the reptiles that they have produced are described in references 1, 4, 9–11.

At the end of the Triassic period, the scenery around Bristol consisted of limestone hills in which deep fissures and caves formed. The climate was subtropical with wet and dry seasons, and the vertebrate fauna consisted of dinosaurs, small lizard-shaped reptiles and early mammals. Occasionally, these animals fell into the fissures, or were washed in by flash floods with mud and sand. The habitat was assumed to have been a rocky 'upland',

Michael Benton is in the University Museum, Parkes Road, Oxford, OX1 3PW.

well away from the lush pastured dinosaur country below. However, there is now strong evidence for marine influence in at least one fissure (Tytherington); marine microfossils (acritarchs and dinoflagellates) occur, kerogen derived from marine algae is abundant¹ and a glauconitic clay mineral has been found which could indicate brackish conditions². This fissure at least must have been located near the coast with its base below sea-level. This would permit mixing of the meteoric waters that washed the debris into the cave with saline waters from the sea.

We must now view the fissure animals as representing a sample of a lowland fauna rather than a peculiar upland fauna. Very large animals are excluded — like the five-metre long dinosaur *Plateosaurus* which is found widely in Europe at the same time. However, many single bones of its smaller relative *Thecodontosaurus* have been found in various fissures. The most abundant animals, however, are lizard-sized, from 10–50 cm in total body length. A detailed study of the ecology of an assemblage of these small animals from Slickstones Quarry has just been published⁴, and it gives us a unique impression of the fauna scuttling around beneath the shadow of the dinosaurs.

The Slickstones fauna consists of 12 or more forms, and includes abundant remains of six different sphenodontids. Most of these probably ate insects and other invertebrates, as does the living *Sphenodon*, but one form at least may have been herbivorous — a conclusion suggested by patterns of tooth wear. The authors go on to suggest that differences in the composition of the 'faunas' found in different fissures were caused by inter-specific competition for food resources⁴. This seems to be stretching the evidence a little far, unless we assume that the beasts actually lived in the caves. The sphenodontids were preyed upon by dinosaurs and other archosaurs which are represented in the fissure fossils by a variety of jaws and isolated sharp knife-like teeth.

The Bristol fissures should produce a great deal more important information. Many of the fissure reptiles have yet to be described and the careful collecting techniques being used should also yield much new information on their palaeoecology. There is a rare satisfaction in finding out about some of the *small* reptiles that have now been shown to have lived during the Age of the Dinosaurs. □

Archaeology

Stonehenge comes alive

From Julian Richards

THE antiquarians of the seventeenth and eighteenth centuries, the barrow diggers of the nineteenth, and the excavators, fieldworkers and astronomers of the twentieth have all been drawn to the Stonehenge area to observe, to dig and to speculate. To many it has long been apparent that the visible remains, barrows, henges and cursus monuments, while evidence of intensive and consistent activity during the Neolithic and Bronze Ages, provided only a partial picture of prehistoric society. Ideology and death left their enduring traces, but if we could see the dead, then where were the living?

'Stonehenge and Its Environs' (Edinburgh University Press) is the title of the survey carried out by the Royal Commission on Historic Monuments which has started to provide clues in the hunt for the living. Traces of now flattened prehistoric fields, small enclosures and boundary ditches point to an organization of the landscape beyond ritual and burial and provide an invaluable framework on which the next phase of investigation could be based. In 1980, the Trust for Wessex Archaeology began a programme of survey in the area of some 2,500 hectares studied by the Royal Commission. The overall brief of this Stonehenge Environs Project, commissioned by the Inspectorate of Ancient Monuments, was the location and evaluation of areas of prehistoric activity and the provision, by sample excavation, of firm indications of preservation for specific monuments.

With the exception of rare pockets of old grassland and woodland, and areas recently returned to grass by the National Trust, the arable landscape immediately dictated the most suitable field survey approach: surface collection 'field-walking'. The systematic collection of artefacts from the weathered surface of ploughed fields is becoming increasingly recognized as a major field survey technique. The use of flint and similar lithic material for tools throughout prehistory provides one enduring component of a total artefact assemblage that may also have included bone, metal and pottery.

The two-stage surface collection strategy used in this and other Trust for Wessex Archaeology projects aims initially to locate widespread patterns and anomalous areas. These are assessed in the light of localized topography and soil changes. Collection is initially by means of transects within hectare squares based on the national grid and provides an initial surface sample of approximately 8–10 per cent. The second stage of more detailed gridded collection tends to be reserved as an intensive pre-excavation sampling technique when it is carried out in con-

junction with geophysical and geochemical surface surveys.

Over the past three winters more than 1,000 hectares have been examined and patterns of prehistoric activity have been recorded. Far from being a 'ritual landscape', devoid of inhabitants, the recent survey has shown areas of intensive domestic and industrial activity contemporary with the major ceremonial centres. Even within the relatively restricted study area, the patterns represented by the surviving monuments can be seen reflected in shifting zones of domestic activity culminating in the developed agricultural and territorial landscape of the later Bronze Age, c. 1000 BC.

Survey and sample excavation have thus provided a spatial, functional and chronological framework for prehistoric activity around Stonehenge, the only missing component being evidence for environmental change. The potential of upland chalk areas to provide environmental evidence is limited largely to land snails within protected contexts, particularly in buried soils and within ditch deposits. The palaeoenvironmental potential of the Stonehenge area, with its barrow mounds, ditches and other potentially dateable earth works, is thus extremely high and has been exploited in tandem with the main project. Research funding from the Society of Antiquaries, the British Academy and the Prehistoric Society has facilitated the investigation and analysis of a number of overlapping dateable environmental sequences.

The field aspects of the project are now almost complete, with only a short season of surface collection and recording of field monuments remaining. The project will then progress to a phase of full-time analytical work intended to produce both management strategies and a full synthetic publication of the projected results.

It is hoped to continue research in the area, particularly on the aspects of the prehistoric environment, but this will now depend on the availability of research funds other than from the UK Department of the Environment.

It is only by moving beyond empirical observations that we can hope to comprehend the societies in which Stonehenge was built. These must be understood if we are to preserve in a sensible manner their tangible, if ephemeral, remains. It is hoped that the Stonehenge Environs Project will not only have added to our understanding of the area, but also provided the means by which it can be both preserved and explained. □

Julian Richards is at the Trust for Wessex Archaeology, The Archaeological Centre, 65, The Close, Salisbury, SP1 2EN.

1. Marshall, J.E.A. & Whiteside, D.I. *Nature* **287**, 627 (1980).
2. Whiteside, D.I. & Robinson, D. *Palaeogeogr., Palaeoclimatol., Palaeoecol.* **41**, 81 (1983).
3. Fraser, N.C. *Palaeontology* **25**, 709 (1982); (in the press).
4. Fraser, N.C. & Walkden, G.M. *Palaeogeogr., Palaeoclimatol., Palaeoecol.* **42**, 341 (1983).
5. Evans, S.E. *Zool. J. Linn. Soc.* **70**, 203 (1980).
6. Evans, S.E. *Zool. J. Linn. Soc.* **73**, 81 (1981).
7. Owen, R. *Monogr. Palaeontogr. Soc.* **24** (1871).
8. Swinton, W.E. *Ann. Mag. nat. Hist.* (11) **4**, 591 (1939).
9. Kermack, K.A., Mussett, F. & Rigney, H.W. *Zool. J. Linn. Soc.* **53**, 87 (1973).
10. Robinson, P.L. *J. Linn. Soc. (Zool.)* **43**, 260 (1957).
11. Halstead, L.B. & Nicoll, P.G. *Stud. Speleol.* **2**, 93 (1971).

Neurogenesis in *Drosophila**Notch* mapping function obscure

from Michael Akam

WHY does the *Notch* locus of *Drosophila melanogaster* attract so much attention, when the phenotype for which it is named is merely the generation of notches in the tips of the wings? The locus has for example now been cloned by two independent groups^{1,3}.

There are good reasons for this interest. Complete lack of function of the *Notch* locus results in a much more drastic developmental defect than notching of the wings — great increase in the proportion of ectodermal cells which become neuroblasts⁴⁻⁷, consequentially an extensive hypertrophy of the nervous system at the expense of the epidermis and, eventually, the death of the embryo.

The notching phenotype is the dominant, but developmentally trivial, effect of heterozygosity for such a null mutation. Moreover, *Notch* is one of the best characterized of the genetically complex loci in *Drosophila*^{5,8-12}. In addition to the dominant alleles, there are many recessive alleles which fail to complement *Notch* null mutations so that by conventional criteria they lie in the *Notch* gene. The phenotypes of these recessive mutations are varied. They include lethal mutations which have a range of lethal phases (*1(1)Notch*), homozygous viable mutations which affect the morphology of the adult eye and bristles (*facet* and *split*) and *notchoid* mutations which, when homozygous, mimic the dominant phenotype of null alleles but have little or no effect on the nervous system. Mutations within any one of these classes generally fail to complement one another, but recessive mutations with different phenotypes often show total complementation. Thus by this test, each phenotypic class defines a different gene.

There are also *Abruptex* mutations. These have a dominant wing phenotype distinguishable from that of the null alleles; some are recessive lethals, some homozygous viable; in *trans* with other *Notch* complex mutations they may either enhance or suppress the *Notch* phenotypes.

Many of these mutations have been mapped by genetic recombination^{8,9,11,12}. Recessive alleles and *Abruptex* mutations are interspersed with the null alleles, ruling out any possibility that the latter are deletions of several functional units. The genetic unit which they define is extremely large — 0.13 map units, or 19 times the genetic length of the xanthine dehydrogenase gene at the *rosy* locus².

The aims of the molecular studies at *Notch* are twofold: to understand the molecular basis of the genetic complexity and to identify the *Notch* products, and their role in neurogenesis.

The isolation and delimitation of the *Notch* complex was much facilitated by the numerous deletion and inversion mutations accumulated during decades of genetic analysis. Artavanis-Tsakonas and his colleagues noted that one deficiency¹¹ juxtaposed the *Notch* locus and a segment of the X chromosome which they had already cloned¹. Thus it was one small step to hop into the complex. Both Artavanis-Tsakonas *et al.*^{1,3} and Kidd *et al.*² report extensive chromosomal walks from this starting point. Both also map the location of chromosome breakpoints associated with *Notch* complex mutations, and regions of the complex homologous to transcripts. Kidd *et al.* report further the extensive analysis of the major *Notch* transcript by two dimensional S1 mapping², and Artavanis-Tsakonas *et al.*³ the isolation of cDNA clones derived from this major transcript.

Their results indicate that the *Notch* locus, like other complex loci in *Drosophila*¹³, contains at least one long transcription unit with multiple introns, in this case spanning 35–40 kb. The genetic complexity reflects the complex organization of this transcription unit, but appears also to depend upon the particular characteristics of those mutations which result from the insertion of transposable elements.

There is a close correspondence between genetic and molecular maps. At *Notch*, one map unit corresponds to approximately 280kb of DNA, and *Notch* null mutations span 34kb. A spliced transcript longer than 10kb (ref. 1) and shorter than 11.5kb (ref. 2) is encoded throughout the region spanned by these *Notch* mutations. Kidd *et al.* have shown that it contains at least 8 introns ranging in size from < 1 to 20 kb, and find some heterogeneity at its 3' ends². One region at the 3' end of this transcript contains a repetitive sequence which is homologous to other RNA species³.

In several cases, mutations in the complex are correlated with the insertion of transposable elements into the locus. Four insertions which have given rise to *Notch* null mutations all lie in or very close to the exons of the major transcript². In contrast, four recessive *facet* alleles and one recessive *Notch* lethal carry insertions within the major intron of this transcript^{2,3}. Kidd and his colleagues speculate that stage or tissue specific transcription of the transposable elements in these mutations may result in the regulated inactivation of the *Notch* locus, and thus in the localized phenotype of the recessive mutations.

Shellenbarger and Mohler¹⁰ suggested in 1975 that the diversity of the *Notch* complex recessive mutations results from the

inactivation of a single *Notch* product at specific stages in development. This model, originally proposed to account for the interactions between temperature sensitive *Notch* mutations and the recessive alleles, accommodates remarkably well the molecular explanation now proposed.

By no means all *Notch* locus mutations result from insertions or gross rearrangements. The majority of both irradiation and chemically induced mutations were not detectably rearranged¹⁻³. Most of these have however been mapped by recombination; using the correlation between genetic and molecular maps established for the insertion mutants, the majority of *Notch* null alleles can be assigned to the major exon at the 3' end of the *Notch* transcription unit^{2,3}. Thus they may be point mutations in a *Notch* protein.

From these data it seems that there is no necessity to postulate a multiplicity of products to explain the genetic complexity of the *Notch* locus. However, alternate splicing or processing pathways may yield a family of products from the major transcription unit, and that closely adjacent transcription units may be related to the function of *Notch*^{1,3}.

A very different molecular approach leads Thorig *et al.*^{14,15} to suggest that the *Notch* complex is a multifunctional locus encoding a group of mitochondrial flavoprotein enzymes — NADH oxidase, NADH dehydrogenase, succinate dehydrogenase and glycerophosphate dehydrogenase. The activities of these enzymes in flies are dependent on the dosage of *Notch* genes, but the data indicating that *Notch* itself encodes these enzymes are not yet compelling. However, as Thorig *et al.* point out, the *rudimentary* locus of *Drosophila* exhibits a range of mutant phenotypes and complex complementation behaviour, and this locus is now known to be the structural gene for a polypeptide carrying three enzymes of pyrimidine biosynthesis¹⁶.

In what sense then is *Notch* a gene controlling neurogenesis? Wright⁵ suggested that the *Notch* product may be a repressor of neurogenesis — in the current parlance of *Drosophila* development, between epidermal and neurogenic pathways. Recent series papers make it clear that the situation is more complex. In a systematic search for mutations with embryonic phenotypes similar to that of *Notch*, Campos-Ortega and his colleagues^{6,7,17,18} have identified six more 'neurogenic loci'. Lack of any one of these gene functions results in essentially the same developmental switch as that seen in *Notch* embryos — in extreme cases, all the cells in a large region of the ectoderm develop as neuroblasts, and no ventral cuticle is formed.

Michael Akam is an MRC Senior fellow in the Department of Genetics, Downing Street, Cambridge, CB2 3EH and a fellow of Kings College, Cambridge.

If indeed *Notch* does encode mitochondrial enzymes, then one factor affecting the determination of neuroblasts in insects may prove to be a critical balance in some metabolic state of the ectodermal cells. Most of the seven 'neurogenic loci' might then be genes whose products affect this balance. It is perhaps relevant that efforts to find the 'real' neural inductor in amphibians led to the conclusion that the ectoderm is finely balanced between epidermal and neural differentiation, and that many non-physiological stimuli could provide the trigger for neural induction¹⁹.

If however, the products of the 'neurogenic genes' do encode such housekeeping products as flavoprotein enzymes, why is it that so many tissues of the *Drosophila* embryo differentiate apparently normally in their absence?

Whether or not the *Notch* gene makes an 'interesting' product, it is interesting to study the effects of mutation on this complex transcription unit. And lurking among the set of neurogenic loci that Campos-Ortega has identified, there may yet be a selector gene which monitors the state or position of each ectodermal cell, and

mediates its developmental commitment. It remains to be seen whether the genetic and molecular analysis of these genes will elucidate their role in neurogenesis. □

1. Artavanis-Tsakonas, S., Muscavitch, M.A.T. & Yedvobnick, B. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1977 (1983).
2. Kidd, S., Lockett, T.J. & Young, M.W. *Cell* **34**, 421 (1983).
3. Artavanis-Tsakonas, S. *et al. Developmental Genetics* (in the press).
4. Poulson, D.F. *Am. Nat.* **79**, 340 (1945).
5. Wright, T.R.F. *Adv. Genet.* **15**, 261 (1970).
6. Jimenez, F. & Campos-Ortega, J.A. *Wilhelm Roux Arch.* **191**, 191 (1982).
7. Lehmann, R., Jimenez, F., Dietrich, U. & Campos-Ortega, J.A. *Wilhelm Roux Arch.* **192**, 62 (1983).
8. Welshons, W.J. *Science* **150**, 1122 (1965).
9. Welshons, W.J. *Genetics* **76**, 775 (1974).
10. Schellenbarger, D.L. & Mohler, J.D. *Genetics* **81**, 143 (1978).
11. Keppy, D.O. & Welshons, W.J. *Chromosoma* **76**, 191 (1980).
12. Foster, G.G. *Genetics* **81**, 99 (1975).
13. North, G. *Nature* **303**, 134 (1983).
14. Thorig, G.E.W., Heinstra, P.W.H. & Scharloo, W. *Mol. Gen. Genet.* **182**, 31 (1981).
15. Thorig, G.E.W., Heinstra, P.W.H. & Scharloo, W. *Genetics* **99**, 65 (1981).
16. Segraves, W.A., Louis, C., Schedl, P., & Jarry, B.P. *Mol. Gen. Genet.* **189**, 34 (1983).
17. Lehman, R., Dietrich, U., Jimenez, F. & Campos-Ortega, J.A. *Wilhelm Roux Arch.* **190**, 226 (1981).
18. Campos-Ortega, J.A. *Wilhelm Roux Arch.* **192**, 317 (1983).
19. Slacke, J.M.W. *From egg to embryo* Ch.3. (Cambridge University Press).

Geochemistry

Chemical inhomogeneity of mantle above 670 km transition

from Don L. Anderson

WHAT chemical composition the mantle has and on what scale mantle convection takes place are related questions which require data from both seismology and petrology to answer. The mantle, which represents about 68 per cent of the Earth's mass, is usually considered, especially by petrologists, to be mostly homogeneous in composition. Discontinuities within it are generally assumed to represent isochemical phase changes in an olivine-rich material rather than chemical boundaries. Thus, the 400 km and 670 km boundaries detected by seismic means are taken to represent the 'olivine-spinel' and 'spinel-post-spinel' phase changes, respectively. The 670 km discontinuity is the more significant, for seismic velocities increase greatly at this depth. This boundary must be very sharp as it is both a good reflector of short-period seismic waves and is the lower boundary of earthquake activity — no reliably-located earthquake has ever been found below it. These observations can be used to argue that the '670' is a chemical interface that provides a barrier to convection.

A.C. Lees, M.S.T. Bukowski and R. Jeanloz (*J. geophys. Res.* **88**, 8145; 1983) have recently computed the reflection coefficient for a variety of plausible phase changes and conclude that mineralogical and seismic data are compatible with a two-layered mantle that has a discontinuous change in both phase and composition at

670 km, and that the change in material properties must occur over a region less than 3 km thick. Phase changes are typically spread out over a much greater depth interval and give reflection coefficients more than an order of magnitude less than those observed. This conclusion is not new but it serves to keep up the pressure on those who advocate that the mantle is chemically uniform and is involved in convection as a whole. Although the nature of the 670 km discontinuity is important, Lees and colleagues' results may produce a yet more significant conclusion.

They calculate the velocity in the transition region for olivine, for orthopyroxene and olivine and for garnet mineral assemblages, including the effects of phase changes. The velocities in all these minerals and assemblages are much greater than the seismic velocities in the transition region and there is no similarity in shape or depth between the mineralogical models and the 400 and 670 km mantle discontinuities. Taken at face value, these results would seem to rule out peridotite, pyrolite or any olivine-orthopyroxene rich assemblage for the 400–670 km depth interval and the olivine-spinel-post-spinel explanations for mantle discontinuities. Although the temperature and temperature gradient can be adjusted to give a fair match of olivine-rich aggregate densities with the seismically-determined density, the much

more accurately determined velocities are poorly matched. The velocity and density gradients of the transition region are greater than calculated for the peridotite minerals. A spread-out phase change or a compositional gradient is implied.

Lees and colleagues do not pursue these discrepancies in the region above 650 km. They restrict themselves to conventional mineralogies involving (Mg,Fe)₂SiO₄, (Mg,Fe)SiO₃ and (Mg,Fe)₃Al₂Si₂O₁₂. They assume that the MgO, FeO and SiO₂ contents of mantle silicates control the physical properties and the locations of phase boundaries, but, as usual, the effects of CaO, Al₂O₃ and Na₂O are ignored. These are minor components of peridotite and pyrolite minerals, the presumed dominant minerals of the mantle. Eclogite, an alternative mineralogy for the transition region, is dominantly clinopyroxene (diopside and jadeite) and garnet, transforming gradually to garnetite (CaO-Al₂O₃-Na₂O-rich garnet solid solution) over the pressure interval of the transition region. Unfortunately, velocities in an eclogite transition region, including phase changes, are not treated, in spite of the obvious discrepancies with the peridotite minerals.

The idea of a peridotite mantle dominated by olivine is deeply entrenched in geology. Olivine and ortho-pyroxene are dominant minerals in xenoliths brought to the surface in alkali basalts and kimberlites and in upthrust segments of the mantle exposed in fold belts and oceanic fracture zones. The most abundant rocks emerging from the mantle, however, are basalts. At high pressure, these rocks are composed mainly of garnet and clinopyroxene — they are eclogites. At one time it was thought that the upper mantle and the source region of basalts were primarily eclogite. More recently, it has been assumed that the basalt-eclogite fraction of the mantle is dispersed and is liberated by 20–30 per cent partial melting, leaving behind a refractory olivine-orthopyroxene-rich residue. The fact that olivine-orthopyroxene aggregates cannot explain the 400 and 670 km discontinuities, and the velocity and velocity gradients in the transition region, will force a reexamination of the basalt source region, the chemical uniformity of the mantle, and the composition of this region of the mantle.

It is also important to look at clinopyroxene-garnet-rich assemblages. Lees *et al.*'s inability to match the compressional velocity is particularly significant because theirs is not an *ab initio* calculation. They used the seismic data themselves to constrain the velocity-bulk modulus ratio. Those who would argue for a chemically uniform mantle and an olivine-rich transition region are now faced with a formidable array of contrary evidence. □

Don L. Anderson is in the Seismological Laboratory, California Institute of Technology, Pasadena, California 91125.

Living plant fossils

SIR — The article "Seeds of thought for plant conservationists" by P.D. Moore¹ is of great interest but it should be added that seeds are not the only plant reproductive parts that remain viable over a considerable period of time in datable contexts. The same is true of fungal mycelia and spores², bacteria^{3,4}, apparently not viruses⁵ but probably algae, which can survive as cysts.

Moore² isolated 61 microfungi from two peat profiles collected in the Dublin-Wicklow mountains, Ireland, using Warcup's soil plate method. More viable remains occurred at the surface but the number of viable taxa (6 or 7) was constant with depth. Unfortunately the peats have not been dated by radiocarbon methods.

Bacteria of the thermophilic actinomycete genus *Thermoactinomyces* are relatively easily isolated and identified⁵. Their resistant endospores survive for longer than 1,500 years in anaerobic lake muds³ and have been cultured using occupation debris from the Roman fort site of Vindolanda, Northumberland⁴. The concentration of colonies grown varied with sediment type. The microbiology of grain stored in Iron Age-type pits has also been studied⁷ as part of an experimental project on agricultural practices.

Conservation should not be confined solely to higher plants and the medical value of *Penicillium* is a good indication why. Soils that are totally anaerobic or clay-rich (holding water tightly in the clay-humus bond) may serve as storehouses of viable remains of lower plants in addition to seed banks, as Moore suggests.

Definition of extinction in lower plants may be even more difficult than for angiosperms, given the possible longevity of spores.

The palaeoecological importance of these findings requires stressing. While little research on fossil Holocene micro-fungal remains has so far been carried out in Europe, interest may grow if it can be demonstrated that data which add to interpretations of vegetation history based on pollen analysis may thereby be acquired. Microfungi are impossible to identify taxonomically using mycelial remains preserved in peats, and accurate determination using the various spore types is very difficult⁸, but if viable remains occur the prospects are much better.

The various spore types of *Ceratocystis ulmi* (Dutch elm disease) are extremely small⁹ and are probably not identifiable by light microscopy. Finds of fossil elm bark beetles would provide no proof that they were carrying the disease. However, if *C. ulmi* spores remained viable for 5,000–5,200 years, the theory that the elm decline in the British Isles was due wholly or partly to this disease could be properly tested. It has already been shown⁶ that *Thermomyces vulgaris* (actinomyces) spores are plentiful in samples from

Seamere, Norfolk, which have cereal and associated weed pollen and *Thermomyces* spp. occur today in association with cereal cultivation and storage.

Finally, it is of note that Seaward *et al.*⁴ found viable endospores of *T. vulgaris* and *T. dichotomica* at Vindolanda differing from those regarded as typical of currently recognized species. So resuscitation can add potentially to genetic diversity.

B.K. MALONEY

Department of Geography and
Palaeoecology Centre,
Queen's University,
Belfast BT7 1NN, Northern Ireland

1. Moore, P.D. *Nature* **303**, 572 (1983).
2. Moore, J.J. *Scient. Proc., R.D.S.*, **26**, 379–415 (1954).
3. Cross, T., Attwell, R.W., in *Spore Research 1973* (eds Barker, A.N., Gould, G.W. & Wolf, J.) (Academic, London, 1974).
4. Seaward, M.R.D., Cross, T. & Unsworth, B.A. *Nature*, **261**, 407–408 (1976).
5. Shackley, M. *Environmental Archaeology* (Allen and Unwin, London, 1981).
6. Unsworth, B.A., Cross, T., Seaward, M.R.D. & Sims, R.E.J. *J. appl. Bact.*, **42**, 45–52 (1977).
7. Lacey, J. *Stored Prod. Res.*, **8**, 151–154 (1972).
8. Van Geel, B. *thesis*, Univ. Amsterdam (1976).
9. McNeel, D.J., Kulkarni R.K. & Nickerson K.W. *Can. J. Bot.*, **61**, 1349–1352 (1983).

Plants from culture

SIR — Recently there has been an increasing amount of research on the inheritance of characters in the progeny of tissue culture-derived plants. Genetic analyses of this type are likely to become even more common as the techniques for generating (somaclonal) genetic variants and transformants via tissue culture become more widely applied to crop species. In discussing results of this type of research, it has come to our attention that there is a potential source of confusion in the nomenclature used.

It is common practice to self-pollinate plants regenerated from culture and to analyse the resulting progeny for the inheritance of novel traits. The progeny of such selfed regenerated plants have been referred to variously as the F₁, F₂, R₁, R₂, RF₁, RF₂, SC₂, S₁ and G₁ generation^{1–9}. It is therefore apparent that a consistent nomenclature is needed to clarify the nature of the material being described.

Ideally any system adopted should be logical, easy to use and be compatible with established plant breeding nomenclature. We believe that these criteria can be met by strict adherence to the use of letters to designate the process, followed by numbers to show how many times the process has been performed. Thus the original regenerated plants are to be referred to as the R₁ generation. It follows that successive cycles of cell culture and plant regeneration, without a sexual cycle, would produce R₂, R₃, etc. generations. If the regenerated plants from the R₁ generation are selfed, their progeny become the R₁S₁ generation. Repeated selfing to event-

ually produce homozygous inbreds would progress through the R₁S₂, R₁S₃, R₁S₄, etc. generations. Back-crossing to non-cultured material would similarly be described as R₁BC₁, R₁BC₂, R₁BC₃, and so on. By using nomenclature that is compatible with current plant breeding conventions, the same system can easily be extended to more complex breeding programmes. For example, using maize inbreds, (Mo17 × B73 R₁S₁) BC₂ would represent the second backcross generation derived from the cross of Mo17 by an S₁ progeny of a plant regenerated from a B73 culture. Although such complex situations would not be met by most researchers, we feel that any system adopted should be able to cope with all eventualities.

N.P. EVERETT

S.M. FLASHMAN

Stauffer Chemical Company,
1200 S 47th Street,
Richmond, California 94804, USA

1. Willems, G.J., Molendijk, L., Ooms, G. & Schilperoord, R.A. *Cell* **24**, 719–727 (1981).
2. Bourgin, J.-P. *Molec. gen. Genet.* **161**, 225–230 (1978).
3. Hibberd, K.A. & Green, C.E. *Proc. natn. Acad. Sci. U.S.A.* **79**, 559–563 (1982).
4. Lörtz, H. & Scowcroft, W.R. *Theor. appl. Genet.* **66**, 67–75 (1983).
5. Edallo, S., Zucchini, C., Perenzin, M. & Salami, F. *Maydica* **26**, 39–56 (1981).
6. Gengenbach, B.G., Connelly, J.A., Pring, D.R. & Conde, M.F. *Theor. appl. Genet.* **59**, 161–167 (1981).
7. Otten, L. *et al. Molec. gen. Genet.* **183**, 209–213 (1981).
8. Straub, J. & Schell, J. *Molec. gen. Genet.* **183**, 209–213 (1981).
9. Beckert, M., Pollacksek, M. & Caenen, M. *Agronomie* **3**, 9–18 (1983).

Species named

SIR — It is surely of fundamental importance in research work in the areas of basic biology and genetics that the organisms being studied should be precisely identified. I find it extraordinary, therefore, that in the interesting article by Bernards *et al.* on the growth of chromosome ends in trypanosomes (*Nature* **303**, 592; 1983), that neither the genus nor the species of the trypanosomes studied was given.

DAVID WALLIKER

Institute of Animal Genetics,
University of Edinburgh,
Edinburgh EH9 3JN, UK

BERNARDS AND BORST REPLY — Our experiments were done with *Trypanosoma brucei brucei*, strain 427. We expect, however, that the growth and contraction of chromosome ends will be a property of chromosomes in all African trypanosome species, and possibly in all organisms. We apologize that this expectation has inadvertently crept into our paper and we hope that other readers have found the missing information in our earlier papers, amply quoted in the *Nature* article.

A. BERNARDS

P. BORST

Antoni van Leeuwenhoekhuis,
Het Nederlands Kankerinstituut,
1066 CX Amsterdam,
The Netherlands

What is the solution?

SIR — The centenary of Sydney Ringer's important work¹ on the ionic requirements for maintenance of viability of living tissues *in vitro* has, astonishingly, been unremarked; it was not even mentioned at the recent 29th International Congress of the International Union of Physiological Sciences in Sydney, Australia. In the same field we are approaching the 75th anniversary of Tyrode's contribution² and have just passed the 50th anniversary of the advance made by Krebs and Henseleit³. The use of Ringer's, Tyrode's and Krebs' solutions in biomedical research must be among the most common techniques of science, perhaps even rivaling Student's *t*-test. I estimate that Krebs' solution alone features in more than 10,000 scientific papers a year.

These solutions, and their better nourished relatives, the tissue culture media, represent a veritable scientific Ganges. They have become holy waters, with all of the associated mysticism, dogma, priesthoods and devout acolytes. Nostrums from the ancient scriptures, frequently mistranslated to include idiosyncratic errors, are being perpetuated unquestioningly through generations of disciples. There have even been attempts at minor deification with the eponyms "Ringer" and "Krebs" being applied, respectively, to almost any such solution and to bicarbonate/CO₂ buffered solutions in particular. It has been forgotten that the tissue sustaining properties of normal saline were known before Sydney Ringer and that bicarbonate/CO₂ buffered salines were in use before Krebs and Henseleit.

But none of these solutions was designed to simulate the natural bathing fluid of tissue cells, the *milieu intérieur* of Claude Bernard. All were based on what was known of blood plasma composition and were adjusted empirically to produce a desired response. This has resulted in the use of solutions which maintain grossly abnormal pH, electrical potential and osmotic equilibrium in the isolated tissues they support. Because such phenomena as cellular excitability, drug-receptor interaction and enzyme activity may be profoundly affected, why do these solutions continue to be used?

Certainly, there are occasional glimmers of awareness, as a few authors each year notice some of the unphysiological characteristics of Krebs', Tyrode's and Ringer's solutions. These continual rediscoveries have resulted in a multiplicity of slightly "modified" solutions. The fact that each of these is capable of maintaining tissues in a reasonably functional state should not be surprising. Living cells have to cope with the fluctuations produced by water intoxication or dehydration, high or low salt intake, labile pH and such ionic imbalances as result from muscle exercise during

marathon running, for example. Nor should it be surprising that some cellular functions are best demonstrated in solutions with defined, even highly abnormal, compositions. But, if our objective is to gain insight into *in vivo* function, we should expect normal cellular function only in cells coping with environmental conditions close to normal.

Is there a fundamentalist conspiracy promoting a scotomatal attitude among biomedical scientists so that even those who are in the forefront of the giga-ohm-laser-dye-coupled-computerized-tomographic-recombinant-DNA style of cellular research continue to use string and sealing wax technology for their "physiological bathing solutions"?

I espouse heresy. In 1969, I devised a solution based quantitatively upon the ionic composition of a representative mammalian interstitial fluid⁴. This I termed synthetic interstitial fluid (SIF) and it has been used with great success in many laboratories, notably in experiments on isolated human and other mammalian muscles. SIF may not be the ultimate in physiologically balanced salt solutions but it certainly surpasses those now being used.

A. H. BRETAG

School of Pharmacy,
South Australian Institute
of Technology,
North Terrace,
Adelaide, South Australia 5000

1. Ringer, S.L., *J. Physiol., Lond.* 4, 222-225 (1883).
2. Tyrode, M.V. *Archs int. Pharmacodyn. Ther.* 20, 205-223 (1910).
3. Krebs, H.A. & Henseleit, K. *Hoppe-Seyler's Z. physiol. Chem.* 210, 33-66 (1932).
4. Bretag, A.H. *Life Sci.* 8, 319-329 (1969).

ITP nomenclature

SIR — The interesting recent letter *Nature* (3 November, 1983 p.67-69) and News and Views article (3 November, p.16-17) on the possible second messenger function of inositol 1,4,5-trisphosphate may presage a surge of interest in the function and metabolism of inositol phosphates. It would therefore be helpful immediately to establish acceptable and unambiguous abbreviations of these compounds. Streb *et al.* did to a large degree achieve this aim; for example, they used Ins 1,4,5P₃ to indicate inositol 1,4,5-trisphosphate. However, they unfortunately chose cyclic IMP as their abbreviation for inositol 1,2-cyclic phosphate: their own logic would demand the abbreviation Ins1,2-cyclicP, and cyclic IMP would normally be taken to mean inosine 3, 5-cyclic phosphate.

Hesketh's News and Views item was less careful: the inositol lipids were referred to by the outdated PI and PI-4,4-bisphosphate terminology, rather than as PtdIns and PtdIns4,5P₂, and Ins1,4,5P₃ appeared as ITP, which is internationally accepted as

an abbreviation for inosine triphosphate. Although the abbreviations in which inositol appears as Ins are more cumbersome, they are unambiguous, and this is a great virtue in a field into which many new workers are moving.

Another flaw in the published material was the incorrect citation of our recent paper (ref. 5 of the News and Views item): the correct citation is *Biochem. J.* 212, 733-747; (1983).

R. H. MICHELL

Department of Biochemistry,
University of Birmingham,
PO Box 363, Birmingham B15 2TT, UK

1. Rubenstein, E., Bonner, W.A., Noyes, H.P. & Brown, G.S. *Nature* 306, 118 (1983).
2. Landstreet, J.D. & Angel, J.R.P. *Nature* 230, 103 (1971).
3. Martin, P.G., Illing, R. & Angel, J.R.P. *Mon. Not. R. astr. Soc.* 159, 191-201 (1972).
4. Korchak, A.A. & Syrovatskii, S.I. *Astr. Zh.* 38, 885-897 (1961); *Soviet Astr. AJ* 5, 678-686 (1962).
5. Legg, M.P.C. & Westfold, K.C. *Astrophys. J.* 154, 499-514 (1968).
6. Epstein, R.I. *Astrophys. J.* 183, 593-610 (1973).
7. Channmugam, G. & Dulk, G.A. *Astrophys. J.* 244, 569-578 (1981).

Familial cancers

SIR — Several reports have recently appeared showing that familial retinoblastoma is due to mutations in each of the two copies of an autosomal gene. In a review of this work in *Nature*¹ your reviewer, like most people writing about the genetics of retinoblastoma, implies that Alfred Knudson was the originator of the idea that certain familial cancers such as retinoblastoma could be due to the combination of an inherited mutation and a somatic mutation. Actually this hypothesis was proposed by Robert de Mars in a symposium organized by the University of Texas M.D. Anderson Hospital², two years before Knudson's first paper on the subject.

Until then, familial retinoblastoma had been classified as due to an autosomal dominant mutation, but de Mars made the following point: "... I suggest that many of the pedigrees that are labelled as autosomal dominants ... could actually be interpreted as autosomal recessives. We must relate the terms dominant and recessive not only to the level of the individual as a whole ... but also to the level of the individual cell, or cells, which are involved. I think many pedigrees are consistent with the notion that one of the parents in these families might be heterozygous for a recessive and that the neoplasms appear as a result of subsequent somatic mutations in which individual cells become homozygous for a recessive neoplasm-causing gene."

JOHN CAIRNS

Department of Cancer Biology,
Harvard School of Public Health,
665 Huntington Avenue,
Boston, Massachusetts 02115, USA

1. Gilbert, F. *Nature* 305, 761-762 (1983).
2. de Mars, R. in *23rd an. Symp. Fundamental Cancer Research*, 1969, 105-106 (Williams and Wilkins, Baltimore, Maryland, 1970).

Aspherical heterogeneity of the mantle from phase velocities of mantle waves

Ichiro Nakanishi & Don L. Anderson

Seismological Laboratory, California Institute of Technology, Pasadena, California 91125, USA

Long-period surface waves are used to map the lateral heterogeneity of the upper 100–600 km of the mantle. There is good correlation of velocity with surface tectonics and heat flow. Convergence regions are generally slow for Love waves and fast for Rayleigh waves. Back arc basins have slower than average shallow mantle. Some island arcs show evidence of fast material at greater depth. Deep-seated slow anomalies underlie the Red Sea–Afar region of north-east Africa, western North American–northern East Pacific Rise, Indian Ocean triple junction and the Tasman Sea–Campbell Plateau regions. The fastest regions are in the north central Australia–New Guinea and the South Atlantic.

SURFACE waves or, equivalently, short-period free oscillations are sensitive to the seismic velocities and density of the upper mantle and provide a powerful method for mapping lateral heterogeneity of this part of the Earth.

We have measured surface wave phase velocities over about 300 paths by applying a single-station method¹ to long-period fundamental-mode Love (G) and Rayleigh (R) waves from 13 large earthquakes. We assume that surface waves propagate along the great circle path connecting an earthquake and a station and that they can be expressed as travelling waves generated from a double-couple point source. In this study we measured apparent phase velocities of G_2 , G_3 , R_2 and R_3 , and thus have information for both long and short arcs.

We expand the measured velocities in spherical harmonics to avoid an *a priori* association with surface tectonics. A motivation of this study and our previous study² is that the commonly used great circle velocity measurement method³, or, equivalently, the free oscillation method⁴ gives only the even-order terms of a spherical harmonic description of the Earth's lateral heterogeneity^{5,6}. This information, of course, is not sufficient to describe adequately the heterogeneity. The single-station observations provide the missing information, although the measurement accuracy is inferior to that of the great circle observations. To suppress the instability due to the poorly determined odd harmonics we apply a filtering procedure to the inverted spherical harmonic coefficients^{7,8}. Finally, we correlate the resulting phase velocity distribution with surface tectonics, heat flow and the geoid. A detailed description of our single-station velocity measurements, filtering procedure and interpretation, will be presented elsewhere (ref. 7 and C. H. Nataf, I.N. and D.L.A., in preparation).

Single-station velocity measurement

We processed long-period fundamental-mode surface waves generated from 13 large shallow earthquakes recorded on long-period digital networks (IDA and GDSN). The transversely rotated horizontal and vertical component seismographs are used for analysing Love and Rayleigh waves, respectively. The number of individual paths (short arc plus long arc) is larger than about 300 for both waves. The great circle paths for Rayleigh waves are shown in Fig. 1*b*. Coverage is less complete for Love waves.

For the single-station method we have to incorporate information about the source process. In this study we correct the source effect by using source mechanisms and source process times obtained by Nakanishi and Kanamori⁹. The directional part of the source finiteness is ignored. Epicentral locations and origin times are from the National Earthquake Information Service.

To calculate the distance along the great circle path we assume a geometric flattening $fg = 1/297.001$ and a mean radius $R_0 =$

6371.211 km. We do not correct for the frequency dependent ellipticity^{10–12}. This correction does not affect the conclusions of the present article⁸.

Geographical variation

We expand the phase slowness $1/C(\Omega, \omega)$ in terms of spherical harmonics:

$$1/C(\Omega, \omega) = \sum_{l=0}^{\infty} \sum_{m=-l}^{m=l} s_{lm} Y_{lm}(\Omega) \quad (1)$$

where $\Omega = (\theta, \phi)$, ω is the angular frequency, and $Y_{lm}(\Omega)$ are the fully-normalized spherical harmonics. We truncate equation (1) at $l = l_{\max}$ and write the observed phase delay along the propagation path

$$t_{ES} = \sum_{l=0}^{l_{\max}} \sum_{m=-l}^{m=l} s_{lm} \int_E^S Y_{lm}(\Omega) ds \quad (2)$$

If the number of observations is much larger than the number of coefficients s_{lm} , and the surface wave paths are well distributed over the Earth's surface, we can determine the coefficients s_{lm} by solving equation (2) in a least-squares sense and can synthesize $C(\Omega, \omega)$ by equation (1).

We processed G_2 , G_3 , R_2 and R_3 . G_2 and R_2 travel the long arc between source and receiver. G_3 and R_3 travel in the opposite direction and travel the short arc and one complete additional orbit. To calculate the path integral of Y_{lm} in equation (2) for these waves, we used Backus' solution⁵ of the great circle integral of spherical harmonics and numerical evaluation of the integral for G_1 and R_1 .

We solved equation (2) and determined s_{lm} . Since $l_{\max} = 8$ inversion does not appreciably improve the fit over the $l_{\max} = 6$ inversion⁷, we discuss the solutions for $l_{\max} = 6$. We could synthesize $C(\Omega, \omega)$ by straightforward application of the sum formula (1). The coefficients of $l = \text{odd}$, however, have larger uncertainties than those of $l = \text{even}$ harmonics. To suppress the contribution of the poorly constrained odd harmonics we apply a filter based on the method of Whaler and Gubbins¹³.

Figure 2 shows the filtered velocity contour maps for Love waves at for 100, 153, 200 and 250 s. The overall pattern of the Love wave phase velocity variations shows a general correlation with surface tectonics. These waves are most sensitive to the S_H velocity in the upper 200 km (153 s) to 300 km (250 s) of the mantle.

The lowest velocity regions are located in the southeastern Pacific, western North America, north-east Africa centred on the Afar, the central Atlantic, the central Indian Ocean, and the marginal seas in the western Pacific. A high-velocity region is located in the north central Pacific, centred near the Hawaiian swell. Love wave phase velocities are low along portions of the

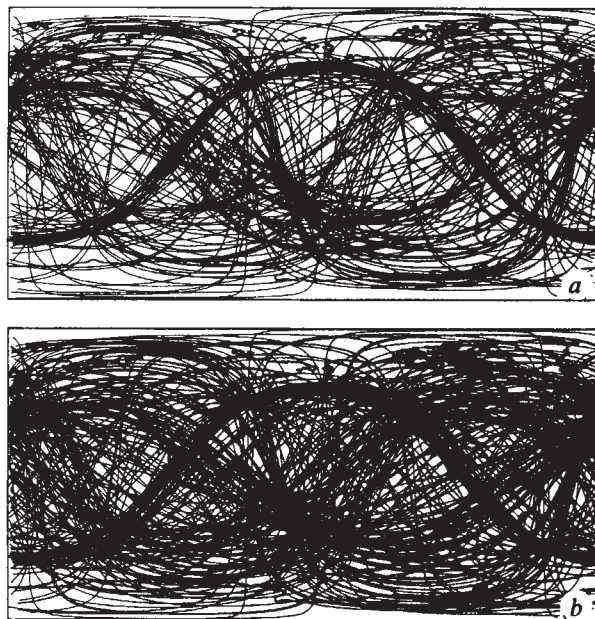


Fig. 1 Great circle paths for Love wave (a), Raleigh wave (b) observations. The same map projection as Figs. 2, 3, 4 and 6.

Mid-Atlantic Ridge, especially near the triple junctions in the north and south Atlantic.

High-velocity regions in continents generally coincide with Precambrian shields and Phanerozoic platforms (northwestern Eurasia, western and southern parts of Africa, eastern parts of North and South America, and Antarctica). Tectonically active regions, such as the Middle East centred on the Red Sea, eastern and southern Eurasia, eastern Australia and western North America are slow as are island arcs, such as the southern Alaskan margin, the Aleutian, Kurile, Japanese, Izu-Bonin, Mariana, Ryukyu, Philippine, Fiji, Tonga, Kermadec and New Zealand arcs.

Rayleigh wave phase velocities are presented in Fig. 3. Rayleigh wave phase velocity also correlates well with surface tectonics, particularly at short periods. The periods considered in this study are most sensitive to S_V velocity between about 100 and 400 km (153 s) to 600 km (250 s) and thus sample deeper than the Love waves.

The slowest regions, for Rayleigh waves, are centred on the Red Sea-Afar, Kerguelen-Indian Ocean triple junction, western North America centred on the Gulf of California, the north-east Atlantic, and Tasman Sea-New Zealand-Campbell Plateau. The fastest regions are the western Pacific, New Guinea-western Australia-eastern Indian Ocean, west Africa, northern Europe and the Southern Atlantic.

One difference between Love waves and Rayleigh waves is evident in the island arcs along the northwestern margins of the Pacific Ocean, such as the Aleutian, Kurite, Japanese, Izu-Bonin-Mariana, Ryuku and Philippine regions. These regions are characterized by the high velocity of the Rayleigh waves. An exception is the Aleutian region for a period of 250 s. In the contour map for 250 s (Fig. 3d) a high-velocity region occurs near the Izu-Bonin-Mariana trenches. The difference between Love and Rayleigh wave results may be explained by the difference in penetration depth. The Love wave results indicate that the shallow mantle, in the vicinity of island arcs and back-arc basins, is slow. Rayleigh waves sample the fast material which has subducted beneath the island arcs. Because of the large wavelengths, the lateral extent of the fast material must be considerable. Anisotropy may also be involved. In regions of convergence and divergence the mantle flow can be expected to be generally vertical. In the centres of convection cells, the flow is mainly horizontal. The sense of anisotropy will change

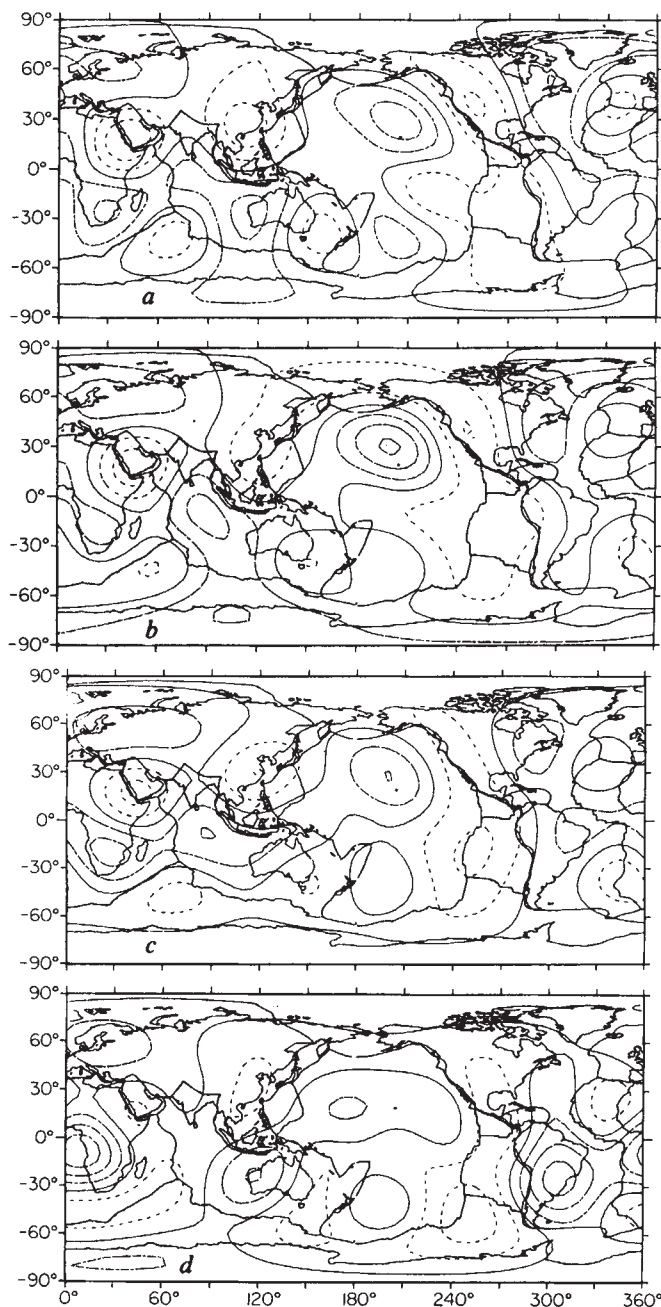


Fig. 2 Love wave phase velocity distributions synthesized by using $l = 1-6$ terms. The solid, chain and broken lines indicate spherical average, higher than average and lower than average velocity. The contour interval is 0.05 km s^{-1} . a, 100 s; b, 153 s; c, 200 s; d, 250 s.

at ridges and subduction zones. This point will be analysed in detail elsewhere.

For Rayleigh waves the Atlantic Ocean is generally high velocity except for the northern part. The low velocities associated with the mid-oceanic ridge system, which appear in the Love wave maps, are not evident in the long-period Rayleigh wave maps. This may indicate that many ridge segments are passive or that $S_V > S_H$ in regions of ascending flow.

Heat flow and geoid

Heat flow and geoid maps have been presented in terms of spherical harmonics and comparison with our surface wave results is straightforward. For this purpose we use the heat flow distribution of Chapman and Pollack¹⁴ and the GEM 8 geoid¹⁵ corrected for the hydrostatic figure of the Earth¹⁶. Chapman and Pollack used observed heat flow supplemented by a tectonic

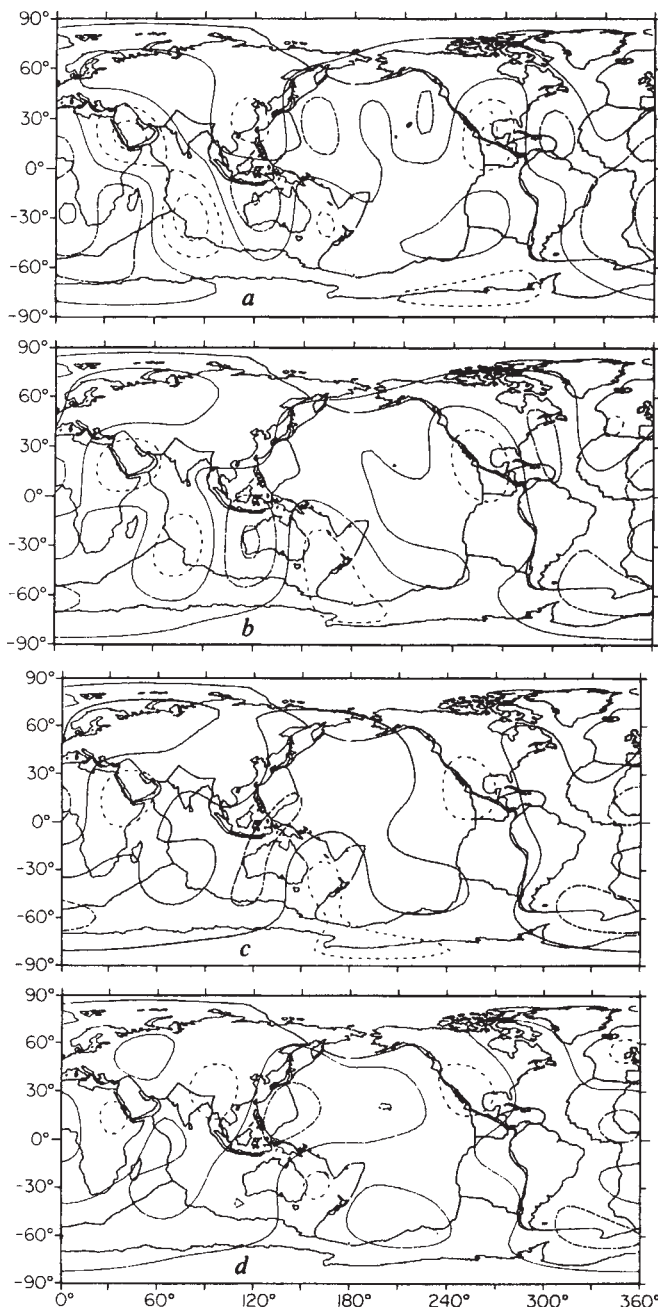


Fig. 3 Rayleigh wave phase velocity distributions synthesized by using $l=1-6$ terms. The symbol conventions are the same as in Fig. 2. *a*, 100 s; *b*, 153 s; *c*, 200 s; *d*, 250 s.

predictor. Figure 4*a* and *b* show the heat flow and geoid maps truncated at $l=6$.

Figures 2-4 show that the phase velocities, especially Love waves, are well correlated with heat flow and, therefore, surface tectonics. The geoid has a weaker correlation with the surface wave velocities.

We compute the degree correlation coefficients γ_l for $l=1$ to 6 by

$$\gamma_l = \frac{\sum_{m=-l}^{m=l} S_{lm} t_{lm}}{\left(\sum_{m=-l}^{m=l} S_{lm}^2 \right)^{1/2} \left(\sum_{m=-l}^{m=l} t_{lm}^2 \right)^{1/2}} \quad (3)$$

where S_{lm} and t_{lm} are coefficients for surface waves and for heat flow or geoid. The results are shown in Fig. 5. The heat flow

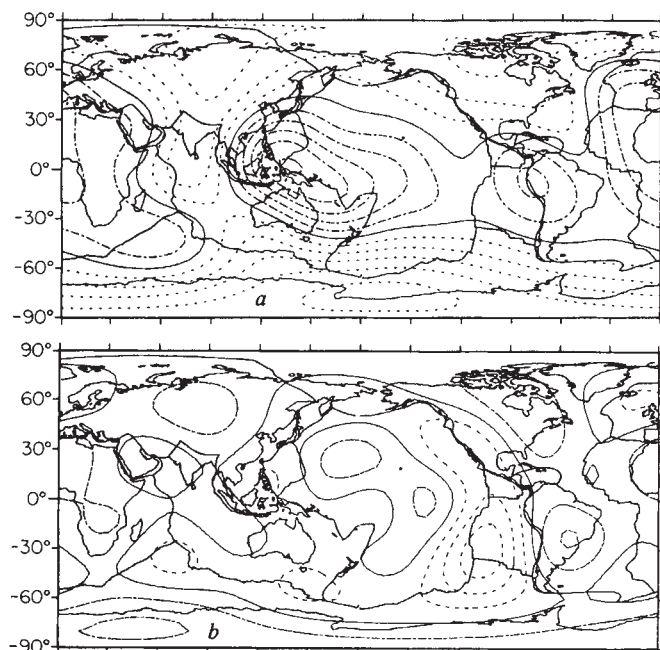


Fig. 4 Contour maps of heat flow and the geoid. *a*, Non-hydrostatic geoid (GEM 8) of Wagner *et al.*¹⁵ and Nakiboglu¹⁶. The contour interval is 20 m. $l=2-6$. The solid, chain and dashed lines indicate the spherical average, higher and lower geoid. *b*, Heat flow of Chapman and Pollack¹⁴. The contour interval is 10 mW m⁻². $l=1-6$. The solid, chain and dashed lines represent the spherical average, lower and higher heat flow, respectively.

has positive correlation with phase slowness for degrees from 1 to 6. The geoid shows negative correlation at $l=2$ (Love and Rayleigh) and 3 (Rayleigh), and positive correlation at $l=4$ to 6.

The correlation of heat flow with Love wave velocities can also be seen by comparing Figs 2 and 4. Because the heat flow map is based on tectonics, as well as heat flow data, this comparison indicates that upper mantle velocities correlate with surface tectonics. Shields are low heat flow areas and exhibit high Love wave velocities, in spite of the thick low-velocity crust. From our Love wave map we would predict that south-east Asia and the Afar region are characterized by more extensive heat flow highs than predicted by Fig. 4*b*. These regions have few heat flow observations. The Canadian Shield and Siberian Platform anomalies are less evident in the Love wave maps than in the heat flow map.

The correlation between surface wave velocities and the geoid is weak. Geoid highs are associated with both subduction and hotspots. Ridges are not evident in either the long wavelength geoid or Rayleigh wave maps. The $l=2$ correlation indicates that the long wavelength part of the geoid has a negative correlation with slowness, that is $l=2$ geoid highs correlate with $l=2$ velocity highs. For shorter wavelength features the reverse is true.

Discussion

Love waves are sensitive to the S_H velocity of the shallow mantle, above about 300 km for the periods considered here. The slowest regions are at plate boundaries, particularly triple junctions. Slow velocities extend around the Pacific plate and include the East Pacific Rise, western North America, Alaska-Aleutian arcs, south-east Asia and the Pacific-Antarctic Rise. Parts of the Mid-Atlantic Rise and Indian Ocean Rise are also slow. The Red Sea-Gulf of Aden-East African Rift (Afar triple junction) is one of the slowest regions. The upper mantle velocity anomaly in this slowly spreading region is as pronounced as under the rapidly spreading East Pacific Rise. Because it also shows up

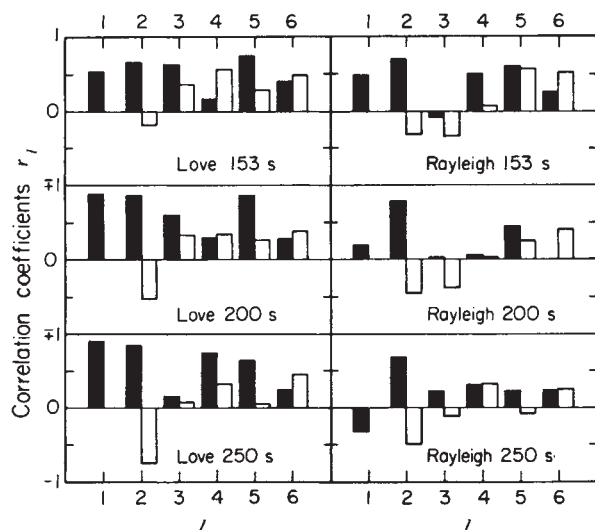


Fig. 5 Degree correlations of surface wave slowness with heat flow and geoid. Solid bars are correlations with heat flow. Open bars are correlations with geoid.

for long-period Rayleigh waves, this is a substantial and deep-seated anomaly. Shields are generally fast, particularly the Brazilian, Australian and South African shields. The fastest oceanic regions are the north-central Pacific, centred near Hawaii, and the eastern Indian Ocean.

Rayleigh waves are sensitive to shallow P-velocities and, for the periods considered here, S_V velocity from about 100 to 600 km. The fastest regions are the western Pacific, western Africa and the South Atlantic. Western North America, the Red Sea area, south-east Asia and the north Atlantic are the slowest regions.

The Precambrian shields are generally faster than average. Silver and Joran¹⁷ distinguish between shields and stable platforms. Okal's regionalization¹⁸, a more suitable model for great circle Rayleigh wave phase velocities^{18,19}, combines these two tectonic regions and has a separate region for island arcs. If we adopt Okal's model, we obtain slower than average Rayleigh wave phase velocity at long period for the stable continental areas^{8,18,19}. In Fig. 3c the velocities in the South Atlantic and the Philippine Sea plate are faster than shields.

Regions of convergence, or descending mantle flow, are relatively fast for long-period Rayleigh waves, particularly near New Guinea, Sumatra, western and northwestern Pacific. These regions have large positive geoid anomalies. Dense, cold material in the upper mantle can explain these results. To explain the surface wave results the fast material must be below the depth sampled by 250-s Love waves and 150-s Rayleigh waves, which is ~300 km depth, and must occupy a considerable volume of the upper mantle. Conductive cooling of the oceanic lithosphere probably extends no deeper than ~150 km. It is doubtful that the subduction of such a thin high-velocity plate could explain the present results and those of our previous regionalization study⁸. Either slabs preferentially subduct in cold regions of the mantle or they pile up in the upper mantle. At shorter periods the velocity of the mantle beneath subduction zones is average or slower than average, presumably reflecting the extensional tectonics and hot shallow mantle under back arc basins. In the tectonic regionalization study of Nakanishi and Anderson⁸, it was also shown that convergent regions are faster than shields at long periods.

It has been proposed that the $l=2$ term is the dominant component of the Earth's asphericity and that it is caused by heterogeneity in the transition region, 420–670 km (ref. 4). Furthermore, it was suggested that the $l=2$ transition zone model was superior to tectonic regionalization. The tectonic model chosen¹⁷, however, ignores convergence regions. A scheme which includes these regions¹⁸ gives results which are

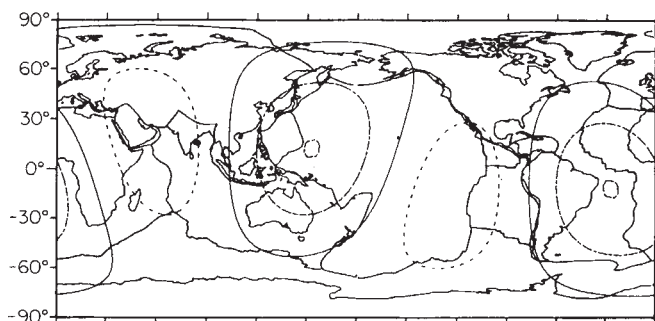


Fig. 6 $l=2$ map of 250-s Rayleigh wave phase velocity. The symbols are the same as in Fig. 2. Rayleigh waves of this period are most sensitive to S_V velocities between depths of 150 and 600 km.

similar to spherical harmonic inversions^{8,20} and show, in addition, that odd harmonics are important. This study shows that there are large lateral variations in the shallow mantle. These plus variations in crustal thickness and water depth are not accounted for in the transition zone model⁴ which attributes all of the observed variation to inhomogeneity of the 420–670 km depth range of the mantle.

A map synthesized from the $l=2$ coefficients at 250 s (ref. 8) is given in Fig. 6. This is characterized by a broad closed high-velocity region centred in the western Pacific, north-east of New Guinea, and a slow velocity region centred on the East Pacific Rise. Because of symmetry the antipodal regions are identical. This map is similar to that presented by Masters *et al.*⁴ but our interpretation is quite different. Because the pattern is similar at short periods^{8,20} it is the shallow mantle as well as the transition region that controls the pattern. In particular, the East Pacific Rise and the Red Sea areas control the location and shape of the lows in the $l=2$ pattern and the fast velocities in the western Pacific–Australia determine the location of the velocity highs. The $l=2$ coefficients are just part of the complete set that is required to describe fully the Earth's asphericity and need no special explanation. The $l=2$ pattern of the associated density field does have special significance because it is involved in the Chandler wobble, tidal friction, the orientation of the Earth's spin axis and polar wander, but again there is no reason to suppose that the $l=2$ pattern requires a special explanation. The tectonic model of Okal¹⁸, for example, based on surface tectonics, has a strong $l=2$ component. The strong $l=2$ component comes from the ocean–continent configuration, the locations of island arcs, and the thermal structure of the sea floor. In particular the velocity highs in Fig. 6 encompass most of the world's old oceanic lithosphere and convergence regions and the lows include most of the young oceanic lithosphere and many of the world's hotspots. The correlation between surface wave velocity and heat flow (that is tectonics) and between velocity and geoid is relatively high for $l=2$, particularly at long periods (Fig. 5). If the rotation axis of the Earth is controlled by density anomalies in the upper mantle, and if density correlates with velocity, then the pattern in Fig. 6 would be symmetric about the Equator.

The most important results of the present study are: (1) discovery of low-velocity anomalies centred on south-east Asia, Afar and the Gulf of California and the demonstration that the latter two persist to long periods, indicating a deep upper mantle origin; (2) the association of high velocities, at long periods, with the western Pacific subduction zones, indicating a large volume of fast material in the upper mantle and; (3) the high velocities in the South Atlantic which do not have an obvious explanation in terms of present plate configurations. The South Atlantic may be underlain by ancient subducted Pacific Ocean lithosphere.

We thank Henri-Claude Nataf, Hiroo Kanamori, and Bradford Hager for suggestions. Jeffrey Given, Fumiko Tajima, and

Jeanne Sauber helped us retrieve the seismograms from the GDSN day tapes at an early stage of this study. The IDA data used in this study were made available by courtesy of the IDA project team at the Institute of Geophysics and Planetary Phys-

ics, University of California, San Diego. This research was supported by NASA grant NSG-7610 and NSF grant EAR811-5236. Contribution no. 3912, Division of Geological and Planetary Sciences, California Institute of Technology.

Received 28 June; accepted 14 September 1983.

1. Brune, J. N., Nafe, J. E. & Oliver, J. E. *J. geophys. Res.* **65**, 247-257 (1960).
2. Nakanishi, I. & Anderson, D. L. *Bull. seism. Soc. Am.* **72**, 1185-1194 (1982).
3. Sato, Y., *Bull. seism. Soc. Am.* **48**, 231-251 (1958).
4. Masters, G., Jordan, T. H., Silver, P. G. & Gilbert, F. *Nature* **298**, 609-613 (1982).
5. Backus, G., *Bull. seism. Soc. Am.* **54**, 571-610 (1964).
6. Zharkov, B. N. & Lyubimov, V. M. *Izv. Phys. Solid Earth* No. 10, 613-618 (1970).
7. Nakanishi, I. & Anderson, D. L. *Geophys. J. R. astr. Soc.* (submitted).
8. Nakanishi, I. & Anderson, D. L., *J. geophys. Res.* (in the press).
9. Nakanishi, I. & Kanamori, H. *Bull. seism. Soc. Am.* (submitted).
10. Dahlen, F. A. *J. geophys. Res.* **80**, 4895-4903 (1975).

11. Dahlen, F. A., *J. geophys. Res.* **81**, 4951-4956 (1976).
12. Dziewonski, A. M. & Sailor, R. V. *J. geophys. Res.* **81**, 4947-4950 (1976).
13. Whaler, K. A. & Gubbins, D. *Geophys. J. R. astr. Soc.* **65**, 645-693 (1981).
14. Chapman, D. S. & Pollack, H. N. *Earth planet. Sci. Lett.* **28**, 23-32 (1975).
15. Wagner, C. A., Lerch, F. J., Brown, J. E. & Richardson, J. A. *J. geophys. Res.* **82**, 901-914 (1977).
16. Nakiboglu, S. M. *Phys. Earth planet. Inter.* **28**, 302-311 (1982).
17. Silver, P. G. & Jordan, T. H. *Geophys. J. R. astr. Soc.* **64**, 605-634 (1981).
18. Okal, E. A. *Geophys. J. R. astr. Soc.* **49**, 357-370 (1977).
19. Souriau, A. & Souriau, M. *Geophys. J. R. astr. Soc.* **73**, 533-551 (1983).
20. Kawakatsu, H. *Geophys. Res. Lett.* **10**, 186-189 (1983).

Frost rings in trees as records of major volcanic eruptions

Valmore C. LaMarche Jr* & Katherine K. Hirschboeck†

* Laboratory of Tree-Ring Research and † Department of Geosciences, University of Arizona, Tucson, Arizona 85721, USA

New data about climatically-effective volcanic eruptions during the past several thousand years may be contained in frost-damage zones in the annual rings of trees. There is good agreement in the timing of frost events and recent eruptions, and the damage can be plausibly linked to climatic effects of stratospheric aerosol veils on hemispheric and global scales. The cataclysmic proto-historic eruption of Santorini (Thera), in the Aegean, is tentatively dated to 1628-26 BC from frost-ring evidence.

VERY occasionally, a major explosive eruption injects large amounts of ash and sulphur aerosols into the stratosphere and upper troposphere. In addition to causing rare atmospheric optical effects—a bluish tint to the Sun and Moon, a milky cast in the daylight sky, and lurid sunsets^{1,2}—such events can have pronounced effects on weather and climate that may persist for several years. Current interest^{3,4} in the climatic effects of volcanic dust and aerosols has been heightened by a series of eruptions of the Mexican volcano El Chichón in 1982 and the eruption of Mt St Helens. A new dimension has been added by the recent development of proxy records of past eruptions from measurements of acidity, chemical impurities and microparticle content of ice cores from the polar regions⁵⁻⁷. We propose here that the occurrence of frost-damage zones in accurately dated tree-ring sequences from subalpine bristlecone pines in the western USA is closely linked to climatological effects of major volcanogenic atmospheric veils, and that these frost rings represent new, independent proxy records of climatically effective eruptions during the past several thousand years.

Frost rings and meteorological events

Frost damage to the wood of mature trees is a rare phenomenon caused by temperatures well below freezing at some time during the growing season, when secondary wall thickening and lignification of immature xylem cells in the annual ring is not yet complete. Freezing promotes extracellular ice formation and dehydration which result in crushing of the outermost zone of weaker cells, leaving a permanent, anatomically distinctive record in the wood⁸. Two successive nights with temperatures reaching -5°C and an intervening day at about the freezing level are sufficient to cause frost damage⁹. Two types of frost rings can be distinguished: earlywood frost damage in the inner part of the ring, which occurs near the beginning of the growing season and latewood frost damage, which typically involves only the dark zone of small, thick-walled cells formed near the end

of the growing season (Fig. 1). If the period of cambial activity is known, observing the position of the damage in a ring often permits dating of the event to within a week or two. If daily meteorological data are available, it is usually possible to identify the specific date on which the damage took place and thus, to characterize the synoptic meteorological situation as well as the antecedent climatic conditions that may have contributed to formation of a frost ring.

Frost-damage zones have been produced in the annual rings of subalpine bristlecone pines (*Pinus longaeva* D. K. Bailey and *P. aristata* Engel.) at intervals of a few decades to a few hundred years for at least the past 4,000 yr. They are observed at localities ranging from California to Colorado, a distance of some 1,300 km. In the course of tree-ring chronology development, the presence and type of frost damage in dated annual rings from living trees¹⁰⁻¹² and sub-fossil wood^{13,14} was routinely noted. It soon became clear that frost damage had not occurred randomly, but that characteristic frost-ring dates were common to many trees at the same site or general locality, and even to localities several hundred kilometres apart. For the period of meteorological record it was possible to identify the dates of damaging related meteorological events because the growing season of bristlecone pine is known fairly well. For example^{15,16}, in the White Mountains, California and the Snake Range, Nevada the frost event that produced latewood damage in AD 1884 (Fig. 1) probably took place on 9-10 September, and similar latewood damage in 1965 probably occurred on 17-19 September. In each case temperatures reached new record lows for the month at nearby meteorological stations. In 1902 an earlywood event recorded in the Snake range almost certainly occurred on 3-4 July when the minimum temperature at upper treeline is estimated to have been -10°C or below. There is clearly a relationship between such unusual meteorological events and the formation of frost rings in subalpine bristlecone pines.

Antecedent climatic conditions may also be important. Study

of two years (1884 and 1965) in which latewood frost damage took place in bristlecone pines in California and Nevada shows that these were notably cool summers in the western Great Basin^{17,18}. One effect of such cool conditions is to delay both the onset and the completion of cambial activity. Bristlecone pines near upper treeline normally begin radial growth about late June, and cell maturation is complete by late August, some 3–4 weeks later than in trees near the lower forest border in the same area¹⁹. During a cool summer, maturation may be delayed to late September, when severely cold weather would

be more likely to occur even during a normal year, and this would widen the tree's 'window of vulnerability' to damaging frost.

Climatic effects of eruptions

One important consequence of a large explosive eruption is the spread of a stratospheric veil of fine silicate ash and sulphur aerosols, with resultant surface cooling²⁰ that may be accentuated at high latitudes by the veil's long residence time in the Arctic stratosphere^{21,22}. The degree of concentration of such a

Table 1 Notable frost-ring events in the western USA and dates of major volcanic events

Frost event		Volcanic event		Estimated DVI/ $E_{\max}$	Combined DVI/ $E_{\max}$
Date	Date(s)	Volcano(es)			
AD 595	—	—	—	—	—
628	—	(see Table 3)	—	—	—
687	—	—	—	—	—
1200	—	—	—	—	—
1227	—	—	—	—	—
1299	—	—	—	—	—
1453	—	(see Table 3)	—	—	—
1485	—	—	—	—	—
(Beginning of Lamb chronology AD 1500)					
No (see Table 3)	1500	Unknown, Java	1,000	1,000	1,000
No	1586	Kelud, Java	1,000	1,000	1,000
No	1593	Ringgit, Java	1,000+	1,000+	1,000+
1601	1601	Unknown	1,000+	1,000+	1,000+
No	1614	Little Sunda Islands	1,000	1,000	1,000
1640	{ 1638 1640 1660	{ Rong, Java Komagatake, Japan Katla, Iceland	{ 500 500 800	1,000	1,000
1660	{ 1660 1660 1660	{ Omate, Peru Pichincha, Ecuador Teon, Banda Sea	{ 1,000 300 300	2,400	2,400
No	1673	Gamma Kunnora, Moluccas	1,000	1,000	1,000
1680	{ 1680 1680 1693	{ Tongkoko, Celebes Krakatoa, Sunda Strait Hekla, Iceland	{ 1,000 400 300	1,400	1,400
No	{ 1693 1694 1694	{ Serua, Moluccas Amboina, Moluccas Celebes	{ 500 250+ 250+	1,300+	1,300+
1732	No	Little Sunda Islands	1,000	1,000	1,000
No	{ 1752 1754 1755	{ Taal, Philippines Katla, Iceland	{ 300 1,200	2,500	2,500
1761	No	Eldeyjar and Jökull, Iceland	2,300	2,300	2,300
No	{ 1783 1783 1785	{ Asama Yama, Japan Vesuvius, Italy	{ 600 2,500	5,400	5,400
1805	No	Mayon, Philippines	300	300	300
1817	{ 1814 1815 1821	{ Tambora, Sunda Strait Eyjafjallajökull	{ 3,000 300	3,300	3,300
No	{ 1822 1822	{ Vesuvius Galunggung	{ 300 500	1,100	1,100
1828	No	Coseguina, Nicaragua	4,000	4,000	4,000
1831	No	Avachinskaya Sopka, Kamchatka	—	4,000+	4,000+
1837	{ 1835 1837	{ Coseguina, Nicaragua Avachinskaya Sopka, Kamchatka	{ 4,000 —	4,000+	4,000+
1866	No	Krakatoa, Sunda Strait	1,000	1,000	1,000
1884	1883	Pelée, Martinique	100	1,000	1,000
1902*	{ 1902 1902 1902	{ Soufrière, St Vincent Santa Maria, Guatemala Katmai, Alaska	{ 300 600 500†	500+	500+
1912	1912	Katmai, Alaska	500†	500+	500+
1941	No	Agung, Bali	800†	800+	800+
1965	1963	Agung, Bali	800†	800+	800+

Notable frost rings are those occurring at two or more localities or in 50% or more of sampled trees in any one locality (years of occurrence in both Great Basin and Rocky Mountains are underlined). Period analysed is AD 560–1969. Major volcanic events are individual eruptions or closely spaced eruption sequences (bracketed) having a combined Dust Veil Index (DVI/ $E_{\max}$) estimated by Lamb¹ as 1,000 or more. Period analysed for volcanic events is AD 1500–1968.

* Earlywood.

† Estimated aerosol injection has been revised upward^{23,27,28}.

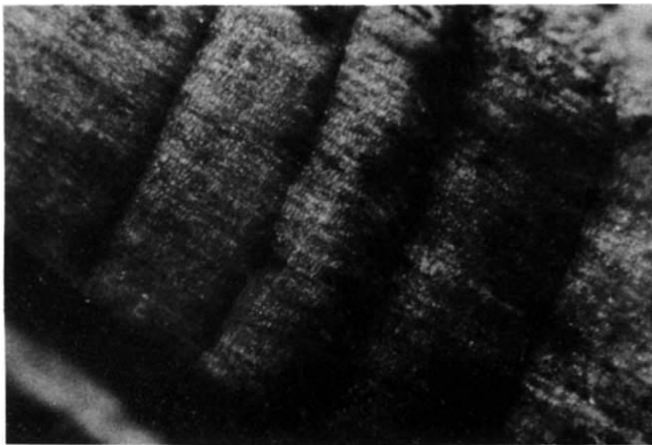


Fig. 1 Latewood frost-damage zone in 1884 annual ring, bristlecone pine, White Mountains, California. Severe frost event occurred 9–10 September 1884.

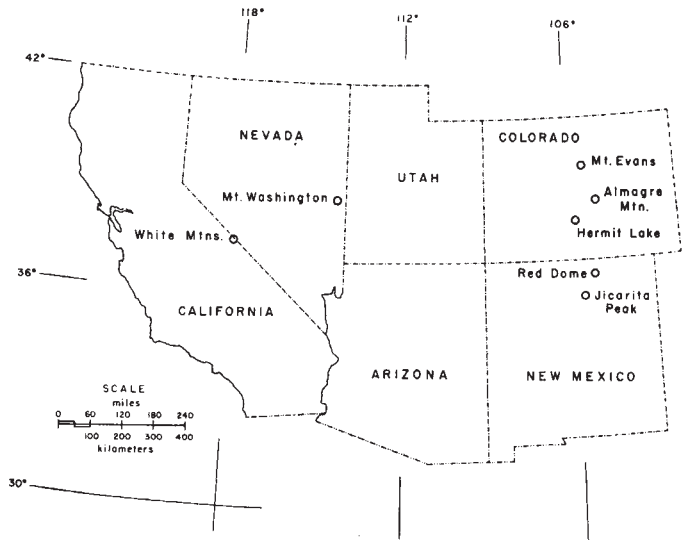


Fig. 2 Location of tree-ring sample localities in the western USA.

cooling effect, and the time lag involved depends in part on the location of the volcano and the composition of the veil²³, on the prevailing circulation, and in part on the season of the year in which the eruption occurs¹.

The effects of an aerosol veil on the large-scale atmospheric circulation have been studied by empirical methods and by modelling. Early work by Wexler²⁴ on the 700 mbar pattern to be expected following a major eruption suggested that a probable circulation response would be expansion of the circumpolar vortex and intensification of an upper level trough over western North America in January. He proposed a summer pattern of westerlies displaced south of normal, and unseasonable synoptic conditions, such that the July weather chart would resemble that of a normal mid-May. Wexler's basic premise seems to be supported by Lamb's observation²¹ of southward displacement of the sub-polar low-pressure zone in the North Atlantic sector

in the first July following a great eruption, and continuing in some cases for 3–4 yr. Flohn²⁵ has speculated that one effect of the intensified polar cold vortices expected to result from high-latitude cooling associated with a stratospheric veil is increased meridionalization of the mid-latitude westerlies. Such quasi-stationary, meridional flow patterns are "... characterized by large-scale standing eddies, extending with diagonal troughs near 200 mbar far into the tropics, and leading to an increased frequency of extreme and unusual weather situations..." Modelling results seem to support the general kind of circulation changes proposed by Wexler, Flohn, and others as consequences of a volcanic aerosol veil. Hunt²⁶, in an early attempt, incorporated volcanic debris in an experiment using a three-dimensional, multilevel general circulation model with a dynamic dust veil of Krakatoa's magnitude. In the short term, he found an overall decrease in zonal wind intensity and changes in the intensity of

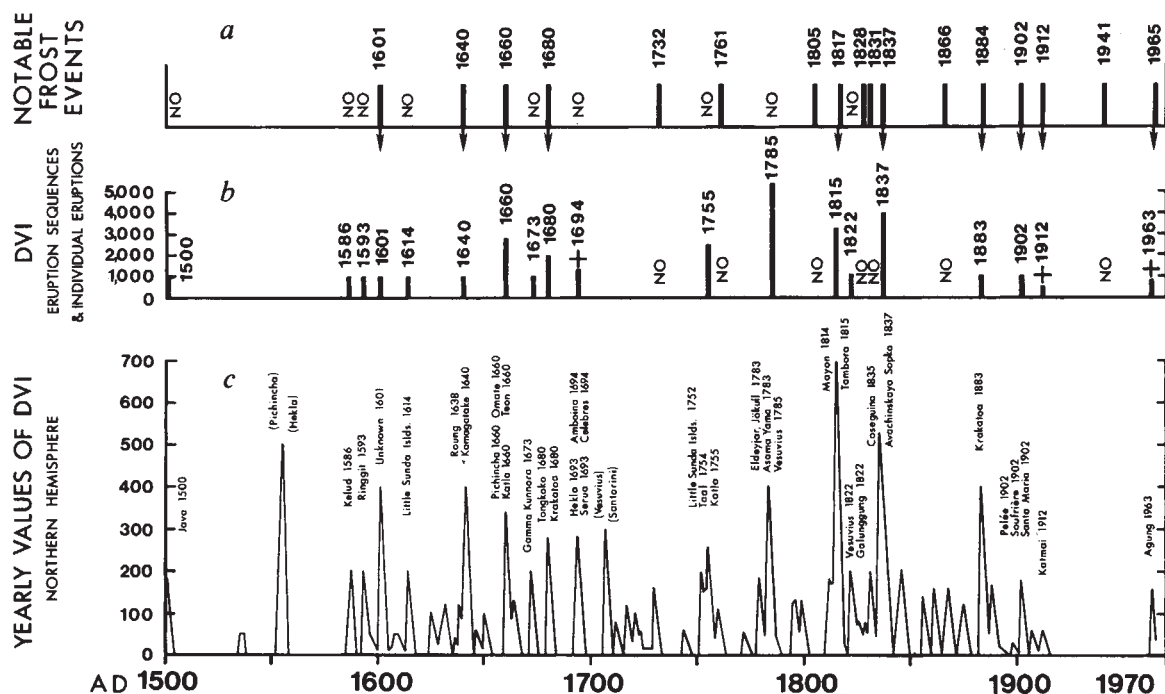


Fig. 3 Frost rings in bristlecone pines in the western USA in relation to major volcanic eruptions on global scale. *a*, Dates of notable frost-ring events. Arrow indicates associated eruption and NO indicates absence of notable frost event at time of major eruption. *b*, Dust Veil Index (DVI/ E_{max}) and dates of eruptions and eruption sequences of 1,000 or greater¹. NO indicates apparent absence of major eruption corresponding to frost-ring date shown in *a*. *c*, Integrated yearly DVI¹, with names of volcanoes and dates of major eruptions. Some other large eruptions are indicated by names of volcanoes in parentheses.

the mean meridional cells, coupled with a decrease in mean hemispheric surface temperature. Cool summers that may contribute to occurrence of frost rings in the western USA seem to be characterized by such increased meridionality, with frequent development of a deep trough at the 700-mbar level in the middle troposphere¹⁰. Individual spring, summer, and autumn months can be 2–4 °C cooler than normal under this synoptic regime. Bearing in mind that the climatic impact of any individual veil will be tempered by such factors as the preeruption state of the Earth–atmosphere system and the location of the volcano, we suggest that the anomalous circulation over western North America in years following great eruptions could consist of southerly displacement of the general westerly flow and/or more frequent development of an upper level trough, accompanied by occasional outbreaks of unseasonably cold air from higher latitudes. Synoptic situations more typical of winter may be expected to occur in late spring and in early autumn. Such a scenario seems to have been followed in the frost-ring year of 1884, where an examination of daily surface-pressure maps for 9 and 10 September shows a very large high-pressure area extending from northern Saskatchewan and Manitoba down through California, Nevada and Utah, probably representing an outbreak of cold Arctic air that took place unusually early, near the end of an already cool and delayed growing season. The mid-May snowstorm and late September cold-wave in 1983 in the Rocky Mountains and High Plains of the western USA could be more recent examples.

Frost-ring records

Data from bristlecone pines at seven localities in the western USA¹² were studied in this work (Fig. 2). These include three sites in New Mexico and southern Colorado, two sites in the Colorado Front Range, one site on the Nevada-Utah border, and four sites in the White Mountains of eastern California. The tree-ring records from the Rocky Mountains begin between AD 560 and 1535. The chronology from the central Great Basin begins in AD 737. Although all of the sites have potential for chronology extension based on records from dead trees and remnants, this approach has been most successful at one site (Campito Mountain) in the White Mountains, where the continuous upper treeline chronology begins in 3435 BC^{11,13,16}. The chronological control provided by these cross-dated tree-ring records representing large numbers of trees ensures accurate placement in time of both the frost-damage event and any associated volcanic eruptions.

Frost rings vary considerably from one event to another in the severity of cell damage, in their frequency of occurrence at a particular site, and in their range of distribution. Branches, seedlings, and very young trees also show a high incidence of frost damage¹⁵, but the results discussed here are based on fairly homogeneous samples of old, well-established trees. With few exceptions, such as 1902, earlywood damage in a given year is restricted to a single locality, perhaps reflecting a local meteorological event or indicating that only at this locality had the trees begun cambial activity, rendering them susceptible to frost damage. However, latewood frost damage is frequently found to have occurred at several localities in the same year, even where these are separated by hundreds of kilometres. For example, several frost events are common to the Rocky Mountains and to the Snake Range, others to the Snake Range and the White Mountains. These variations in severity and geographical extent of latewood damage provided us with a basis for the stratification of frost-ring years that we have used to identify the more important occurrences, termed 'notable frost-ring events'. The results for all of the western USA for the period represented by chronologies from two or more localities are given in Table 1. There are 25 of these notable events in a total of some 116 individual years during which frost damage occurred in at least one sampled tree somewhere in the region. A similar stratification based on less stringent criteria was used

for the White Mountains record alone, covering the much longer period 3435 BC to AD 1971.

Krakatoa effect

The postulated linkage between atmospheric veil effects caused by major volcanic eruptions and the climatological and meteorological setting for severe and widespread frost damage was originally suggested by the remarkable coincidence of frost-ring dates which fell no more than 2 yr after each of the four climatically effective Northern Hemisphere or equatorial eruptions and eruption sequences of the past 100 yr. These dates are 1884 (Krakatoa, 1883), 1902 (Pelée, Soufrière, early 1902), 1912 (Katmai (Novarupta), early 1912), and 1965 (Agung, 1963). In addition to their measured effects on the intensity of the direct solar beam²¹, the aerosol veils associated with most of these eruptions seem to have caused widespread surface cooling^{20,23}. To provide a much longer, if less accurate data set for further evaluation, we referred to Lamb's volcanic eruption chronology¹ and to his dust-veil estimates. A better eruption catalogue is now available²⁸, which is longer, more complete, and probably more objective than Lamb's in its assessment of relative magnitudes, because it incorporates the Volcanic Explosivity Index (VEI) of Newhall and Self²⁹, which emphasizes the explosive eruptions that are most effective in injecting gas and ash into the stratosphere. However, this index does not weight eruptions by geographical location to provide a scaling of probable climatic impact, so we chose to use Lamb's geographically weighted Dust Veil Index (DVI/ $E_{\max}$) in this initial comparison. We identified as 'major volcanic events' those eruptions and/or closely spaced eruption sequences, for which Lamb had assigned DVIs of 1,000 or more. For reference purposes, note that Lamb had scaled his indices to give Krakatoa, 1883, an index of 1,000. There are 19 such events in the total period, AD 1500–1968, covered by his chronology (Table 1 and Fig. 3). In 10 cases, a notable frost event as previously defined occurred in the western USA in the same year or within 1 or 2 yr following a major eruption or eruption sequence. It is also true that in nine cases a frost event occurred without an apparent antecedent volcanic event, and seven volcanic events have no known associated frost event.

To test the statistical significance of this apparent relationship, we separated the total period of concurrent frost and volcanic events into two parts. That period from 1882 to 1968 includes the four well-documented volcanic events cited above and, moreover, is a period for which meteorological data are adequate to characterize climatically-effective eruptions through their depression of hemispheric temperatures as well as to study associated regional climatic and meteorological anomalies. Our approach was to divide the 87-yr period into 29 discrete 3-yr intervals (2–3 yr is a typical decay period for stratospheric aerosol veils²⁰), or triads, and then to count the number of cases in which a notable frost event occurred in the same triad as a major volcanic event. The starting date of 1882 was chosen so that 1883–84 and 1963–65 would be respectively grouped in the same 3-yr intervals, and so that the last year of Lamb's chronology would be included. From elementary probability theory³⁰, the expected frequency of joint occurrences of events in two random, completely independent series is equal to the product of their individual probabilities. Thus, if we assume that these probabilities are equal to the observed average frequencies, we would expect such joint occurrences to take place with a frequency of only 0.024, less than one time in 29 triads (Table 2). The observed number is nearly six times that expected by chance (note here that neither in this nor in the following analysis does the frost-ring date precede that of the associated volcanic eruption within a triad). A similar analysis was performed for AD 1500–1880 that included 127 triads and six joint occurrences. The observed number is over four times the expected chance value. The final step in our evaluation was a contingency analysis using the χ^2 statistic calculated from the observed and expected frequencies of joint occurrence³¹.

Despite the small sample size, the results suggest that for each subperiod, the observed number of joint occurrences of volcanic and frost events is very unlikely to have arisen by chance. Therefore, we can reject the null hypothesis of random association, and accept the alternative hypothesis—that major eruptions are likely to be closely followed by notable frost events—at better than the 99.9% confidence level.

The analytical results show that we can tentatively associate many notable frost events in the western USA over the past several hundred years with antecedent volcanic eruptions catalogued by Lamb. As shown in Table 3 there are other records that can be used to extend this comparison qualitatively much farther into the past, using the long tree-ring record of frost events from the White Mountains, California. However, because it is based largely on data from a single tree-ring site, this frost-event chronology offers a less complete basis for comparison with volcanic events than does the broadly based regional data set given in Table 1. There are also greater uncertainties in the volcanological records in the earlier time periods. We have used the eruption catalogue of Simkin *et al.*²⁸, which is much longer than Lamb's. The initial Greenland ice core data⁵ also provide a long, but locally biased and perhaps less well dated and possibly less complete eruption record spanning

the past several thousand years, based on acidity peaks attributed to volcanogenic aerosols. Finally, we have referred to well-dated eruptions in the late BC-early AD period that are based on careful evaluation² of historical and archaeological records of volcanism and atmospheric veils in the Mediterranean region.

The apparent correspondence of frost events and eruptions is not as good as in the more recent past, but some coincidences are worth noting. Severe frost damage occurred in AD 1601, and may be linked to the acidity peak described⁵ as one of the strongest of the earlier signals in the Greenland ice record. The ice-core data do not reflect a peak at the time of the 1500 frost ring, but an eruption occurred that produced atmospheric optical effects in Europe and is reported by Lamb to have taken place in Java¹. A large eruption of Mt St Helens also occurred about this time²⁸. The severe frost event of 1453 could be linked to a low acidity peak about 1453⁵, possibly related to a large eruption of Kelud in 1451²⁸. The high acidity peak⁵ at about 1258 is not represented in the frost record and, there seem to have been no great eruptions that might account for the numerous frost events from 687 to 1200. However, another strong signal in the Greenland ice is dated at about 623 and may reflect the eruption whose atmospheric veil was recorded in Europe in 626² and which might be related to the 628 frost event. The volcano is unknown, but could be Rabaul, in New Britain²⁸. There is no documentary record of an atmospheric veil at the time of the frost event in 601, and no corresponding acidity peak. However, a major eruption took place in the White River field, Alaska, at about that time²⁸.

Referring to the earliest part of our record, only four notable frost-damage events are recognized in the White Mountains in the 4036-yr period before AD 601. One of these events, in AD 119, falls in a gap between the Crête and Camp Century ice records, but could be related to a major eruption of Ilopango, El Salvador, in the Second or Third Centuries²⁸. Another, dated to 42 BC, might correspond to one of three peaks in the Camp Century core, approximately dated to 260, 210, and 50 BC, respectively⁵. It follows closely the great eruption of Etna, well dated to 44 BC from direct observations². A third frost event, in 2035 BC, is the most severe in the entire record, as it occurs in all the trees in the sample and caused severe anatomical damage¹⁶. There does not appear to be a corresponding acidity

Table 2 Basic data and results of contingency analysis of frost-ring occurrence following major eruptions, AD 1500–1968

Period	1882–1968	1500–1880
No. of triads	29	127
No. of volcanic events	4	15
Observed frequency per triad	0.14	0.12
No. of frost events	5	12
Observed frequency per triad	0.17	0.09
Expected frequency of joint occurrences	0.024	0.011
Expected no. of joint occurrences	0.7	1.4
Observed no. of joint occurrences	4	6
χ^2	16.0	14.7
p	<0.0005	<0.0005

Volcanic and frost events are those listed in Table 1.

* Includes Yates correction for continuity due to small sample size³¹.

Table 3 Notable frost-ring events in the White Mountains, California, and dates of possible associated eruptions

Frost events	Volcano	Volcanic events				VEI
		Lamb ¹	Stothers & Rampino ²	Hammer <i>et al.</i> ⁵	Simkin <i>et al.</i> ²⁸	
<u>2035 BC</u>	St Helens			—	1900?BC	5
<u>1626</u>	Santorini			1390 ± 50 BC	1470 ± 20	6
<u>42</u>	Etna		44 BC	50 ± 20	44	3+
AD 119	Ilopango		—	gap	AD 260 ± 100	6?
<u>601</u>	White River		—	—	525 ± 100	6
<u>628</u>	Rabaul?		AD 626	AD 622–623	540 ± 100	6
687	—			—	—	—
1003	—			—	—	—
1029	—			—	—	—
1077*	—			—	—	—
1099	—			—	—	—
1171	—			—	—	—
<u>1200</u>	—			—	—	—
<u>1453</u>	Kelud?	—		1453–1454	1451	3+
1500	St Helens	—		—	1500	5
	Unknown (Java?)	1500		—	—	—
<u>1601</u>	Unknown	—		1600–1601	—	—
<u>1884</u>	Krakatoa	1883		1883	1883	6

Notable frost rings occur at two or more sites, or in 20% or more of trees (years of occurrence in 50% or more of trees on any site are underlined). Period analysed is 3435 BC to AD 1971. Northern Hemisphere and equatorial eruptions were listed if they have a Volcanic Explosivity Index (VEI²⁹) of 5 or greater²⁸, or were represented by ice core acidity peaks⁵ or well documented atmospheric veils^{1,2}. Source 1 covers the period from AD 1500; 2 from 700 BC to AD 630.

* Earlywood.

peak in Greenland, but the frost-ring date coincides approximately with a large radiocarbon-dated eruption of Mt St Helens²⁸.

Santorini connection

The remaining frost event in the BC-early AD time period, dated to 1626 BC, is especially notable, not only for its severity, but because it offers the intriguing possibility of dating precisely the cataclysmic eruption on Santorini (Thera) in the Aegean Sea in the second millennium BC. The frost-ring date had seemed to place the eruption too early in time to match the conventionally accepted date of 1500–1450 BC based on the archaeological evidence available in the early 1970s³², but this discrepancy in dates may now be resolved.

Because the volume of ejecta has been estimated at several times that of Krakatoa, the Santorini eruption seems likely to have produced a major, climatically effective aerosol veil, which is also likely, on present evidence, to have been linked with a notable frost-ring event. Betancourt and Weinstein³³ have carefully reviewed both the archaeological and radiocarbon evidence for the dating of the beginning of the late Bronze Age in the Aegean. In particular, they focus on radiocarbon dates from material obtained during the excavation of a late Minoan settlement near the village of Akrotiri. It lies on the south side of the largest remnant of the original volcanic island, most of which was either ejected or collapsed into the central caldera following the eruption. The excavations were begun by Marinatos in 1967, and the evidence indicates that the buildings were first toppled by earthquake, and soon thereafter buried under many metres of volcanic ash from the eruption. Betancourt and Weinstein evaluate two sets of dates associated with the destruction of ancient Akrotiri, after first calibrating or correcting the raw dates for atmospheric radiocarbon variability, using the 1973 MASCA tables. [For a fuller documentation and a discussion of these samples see Michael³⁴, who supported the seventeenth century BC eruption date.] The first set of dates is from 'long-lived' material, mostly wood from olive, pine, and cedar that would be expected to predate the eruption considerably; the dates are older than 1700 BC. More relevant here are the dates on six short-lived samples, including charred seeds of legumes from a storage jar, and charred fragments of shrubs. The average date given for these samples is 1688 BC $\pm$ 57 yr. More recently, four additional dates have been reported on grain from storage jars from the destruction level³⁵. They have a large range, but yield an average date of 1675 BC. Given the uncertainties in the individual radiocarbon determinations coupled with the uncertainty in the calibrations themselves, our proposed eruption date of 1626 BC or perhaps 1 or 2 yr earlier, derived from

frost-ring evidence, falls well within the probable range for the true date of the eruption based on radiocarbon.

Conclusions

The potential importance of frost rings as proxy indicators of past eruptions warrants careful weighing of the physical evidence and of the strength of the postulated linkage between aerosol veils, the response of the circulation, and their relation to the climate and weather of the study region. The tree-ring data base is adequate for the past several hundred years, but frost events may be greatly under-represented before about AD 600, because of the decreasing size and geographical extent of the sample. Our method of stratifying 'notable' frost events seems valid, but the summing of magnitudes of closely spaced eruptions (Table 1) may not be justifiable physically in all cases²⁰. The comparison of dates of eruptions and frost rings shows that there is not a one-to-one correspondence. Frost events that are unrelated to aerosol veils have occurred, for example, in 1941. Conversely, there were notable eruptions with no associated frost ring in the western USA. That is, some eruptions may result in a frost record at a certain locality while others may not, and frost damage can be due to atmospheric variability from other causes. We also emphasize that comparisons of the kind that we have attempted become increasingly unreliable in the earlier part of the record. This is due partly to uneven reportage of eruptions and partly to the uncertainties in dating of geological, archaeological, glaciological and other sources of evidence.

Further testing of our hypothesis might include the analysis of an improved frost-event chronology for the western USA using an expanded tree-ring data base. Because frost rings are found in trees in the sub-Arctic and in the high mountains of lower latitudes in many parts of the world^{36–40}, other frost-event records from suitable regions could be compared independently with volcanic chronologies. Our preliminary findings also suggest the need for empirical studies of regional climatic variability and analysis of modelling results to see how important volcanogenic aerosol veils may be as a causal factor in frost-ring formation.

We thank the NSF for financial support of much of the original work on which this research is based, J. M. Mitchell and H. Flohn for stimulating discussions on aerosol veil climatology and P. I. Kuniholm for an introduction to the fascinating but often vexing problems of the late Bronze Age in the Aegean and V. A. S. McCord and A. Allen for help in preparation of the manuscript. We thank W. J. Robinson for permission to reproduce Fig. 2, which originally appeared in the *Tree-Ring Bulletin*.

Received 20 June; accepted 6 October 1983.

- Lamb, H. H. *Phil. Trans. R. Soc. A* **226**, 425–533 (1970).
- Stothers, R. B. & Rampino, M. R. *J. geophys. Res.* **88**, 6357–6371 (1983).
- Kerr, R. A. *Science* **219**, 157 (1983).
- Newell, R. E. & Deepak, A. (eds) *Mount St Helens Eruptions of 1980—Atmospheric Effects and Potential Climatic Impact* (NASA SP-458, US Government Printing Office, 1982).
- Hammer, C. U., Clausen, H. B. & Dansgaard, W. *Nature* **288**, 230–235 (1980).
- Herron, M. M. *J. geophys. Res.* **87**, 3052–3060 (1982).
- Mosley-Thompson, E. & Thompson, L. G. *Quat. Res.* **17**, 1–13 (1982).
- Glerum, C. & Farrar, J. L. *Can. J. Bot.* **44**, 879–886 (1966).
- Glock, W. S. & Reed, E. L. *Science* **91**, 98–99 (1940).
- LaMarche, V. C. *Jr Science* **183**, 1043–1048 (1974).
- LaMarche, V. C. *Jr Nature* **276**, 334–338 (1978).
- LaMarche, V. C. *Jr & Stockton, C. W. Tree-Ring Bull.* **34**, 21–45 (1974).
- LaMarche, V. C. *Jr Quat. Res.* **3**, 632–660 (1973).
- LaMarche, V. C. *Jr & Mooney, H. A. Arct. Alp. Res.* **4**, 61–72 (1972).
- LaMarche, V. C. *Jr in Tree-Ring Analysis with Special Reference to Northwest America* (eds Smith, J. H. G. & Worral, J.) (University of British Columbia Faculty of Forestry Bull. 7, 1971).
- LaMarche, V. C. *Jr & Harlan, T. P. J. geophys. Res.* **78**, 8849–8858 (1973).
- US War Department, Signal Office *Mon. Weath. Rev.* **12** (1884).
- US Department of Commerce *Climatological Data: Utah* (1966).
- Fritts, H. C. *Papers Laboratory of Tree-Ring Research*, 4, (University of Arizona Press, Tucson, 1969).
- Self, S., Rampino, M. R. & Barbera, J. J. *J. Volcanol. Geotherm. Res.* **11**, 41–60 (1981).
- Lamb, H. H. *Climate: Present, Past and Future* Vol. 1 (Methuen, London, 1972).
- Flohn, H. in *Climatic Change and Variability* (eds Pittock, A. B., Frakes, L. A., Jensen, D., Peterson, J. A. & Zillman, J. W.) (Cambridge University Press, 1978).
- Rampino, M. R. & Self, S. *Quat. Res.* **18**, 127–143 (1982).
- Wexler, H. *Tellus* **8**, 480–494 (1956).
- Flohn, H. in *The Physical Basis of Climate and Climate Modelling*, (Global Atmos. Res. Prog. Pub. Ser. 16, 1975).
- Hunt, B. G. in *Climatic Change and Variability* (eds Pittock, A. B., Frakes, L. A., Jensen, D., Peterson, J. A. & Zillman, J. W.) (Cambridge University Press, 1978).
- Mitchell, J. M. *Weatherwise* **35**, 252–259 (1982).
- Simkin, T., Siebert, L., McClelland, L., Bridge, D., Newhall, C. & Latter, J. H. *Volcanoes of the World* (Hutchinson Ross, Stroudsburg, 1981).
- Newhall, C. G. & Self, S. *J. geophys. Res.* **87**, 1231–1238 (1982).
- Mosimann, J. E. *Elementary Probability for the Biological Sciences* (Appleton-Century-Crofts, New York, 1968).
- Blalock, H. M. *Jr Social Statistics* (McGraw-Hill, New York, 1972).
- Lamb, H. H. *Climate: Present, Past and Future* Vol. 2 (Methuen, London, 1977).
- Betancourt, P. P. & Weinstein, G. A. *Am. J. Archaeol.* **80**, 329–348 (1976).
- Michael, H. N. in *Proc. Temple University Aegean Symp.* (ed Betancourt, P. P.) (Temple University Press, 1976).
- Michael, H. N. in *Thera and the Aegean World* Vol. 1 (ed. Doumas, C.) 791–795 (Aris & Phillips, London, 1978).
- Bailey, I. W. *Bot. Gaz.* **80**, 93–101 (1925).
- LaMarche, V. C. *Jr & Fritts, H. C. Z. Gletscher. Glazialgeol.* **7**, 125–131 (1971).
- Dunwiddie, P. W. & LaMarche, V. C. *Jr Aust. For.* **43**, 124–135 (1980).
- Morrow, P. A. & LaMarche, V. C. *Jr Science* **201**, 1244–1246 (1978).
- Dunwiddie, P. W. & LaMarche, V. C. *Jr Nature* **286**, 796–797 (1980).

DNA sequence of the maize transposable element *Dissociation*

H. P. Döring, E. Tillmann & P. Starlinger

Institut für Genetik, Universität Köln, 5000 Köln 41, FRG

The DNA sequence of the terminal 4.2 kilobases (kb) of the 30-kb insertion in the endosperm sucrose synthase gene of maize mutant sh-m5933 shows that it comprises two identical 2,040-base pair (bp) segments, one inserted in the reverse direction into the other. We suggest that the 2,040-bp sequence is an example of the transposable element Dissociation described by Barbara McClintock.

TRANSPOSABLE elements were first detected and characterized in considerable detail by McClintock^{1,2}. Nearly 20 years later, similar elements were studied at the DNA level first in bacteria³⁻⁵ and later in a variety of eukaryotic organisms, notably *Drosophila melanogaster* and *Saccharomyces cerevisiae*^{6,7}. The structure of these latter elements has been elucidated rapidly, including their complete DNA sequences. It would be desirable to study isolated maize transposable elements because of the variety of genetic phenomena associated with them^{1,2,8}. Even when such elements are inserted at the same gene locus, the resulting phenotypes can be quite different. Sometimes, the phenotype resulting from a particular element remaining at the same site is variable in a clonally heritable way ('change in state')^{9,10}. It would be interesting to study the structural correlate of these functional changes. DNA probes derived from the transposable elements might be used to isolate genes previously tagged by the insertion of the element. Eventually, the elements might be used as a vector for the introduction of new genes into chromosomes, as has been demonstrated first in bacteria^{11,12} and recently in *Drosophila*^{13,14}.

As no gene products of transposable elements in plants are known, conventional means of isolating these elements cannot be used. It is possible, however, to isolate DNA clones of genes that have previously been mutated by the insertion of a particular transposable element. Isolation of DNA from wild type and mutant genes reveals the presence of DNA foreign to the gene in question, which can then be identified as the DNA of the transposable element¹⁵. The first plant gene to be used in this way was that coding for endosperm sucrose synthase at the *Shrunken* locus¹⁶. Mutants were available that were caused by the insertion of the transposable element *Dissociation* (*Ds*)^{17,18}. This element cannot transpose by itself, but it can be transposed in the presence of an additional, transposable element called *Activator* (*Ac*) which acts in *trans* on *Ds*². On isolation of genomic clones from *Sh*^{19,20}, two *Ds*-caused mutations were found (ref. 20 and E. Weck, U. Courage-Tebbe, H.-P.D. and P.S., in preparation).

In one of these mutants, *sh-m5933*, the mutation was shown to be due to a complicated event, leading to the insertion of a 30-kilobase (kb) DNA segment, and a duplication of part of this genetic region was found²¹. The DNA contained in the clone was characterized by restriction analysis, and the DNA adjacent to the cloned segment was studied by Southern blot analysis. From these studies, we hypothesized that only the parts of the insertion located at both of its junctions with *Shrunken* DNA, are transposable elements. The DNA located between the two regions of transposable elements is of unknown origin; part of it is unique, other sequences comprise repeated DNA²⁰ (K. Theres, unpublished observations). To test this hypothesis, we sequenced 4.2 kb of DNA adjacent to the 3' end of the *Shrunken* gene. We show here that this sequence corresponds to two identical copies of a 2-kb segment of DNA. One of these segments is inserted in the reverse direction into the other. We

discuss evidence that these 2-kb DNA segments correspond to the genetically defined transposable element *Ds* described by McClintock^{1,2}.

Transposable element insertion

The first 4.2 kb adjacent to the junction between the 3' end of the *Shrunken* gene and the large insertion mentioned above were sequenced. Figure 1 shows the sequencing scheme and the sequence itself is given in Fig. 2. The 4.2 kb sequence is made up of two identical sequences of 2,040 bp. These sequences are found, however, in an unusual arrangement, one being inserted into the other. The two sequences are inverted with respect to each other (see Fig. 3).

We have reported previously that two pairs of inverted sequences are present near the junction between wild type sequences and the insert²⁰. The DNA sequence analysis presented here shows that this really is the case, and also shows that the two pairs of inverted repeats described previously²⁰ are adjacent to each other.

We can ask whether the two 2-kb sequences are transposable elements, and whether the 4-kb structure has been formed by the insertion of one such element into a copy of itself. Many transposable elements terminate in inverted repeats, and almost all of them are flanked by short direct repeats of the recipient sequence. If these structures are found at the termini of a DNA sequence, this sequence may be a transposable element. Figure 2 shows that by these criteria, the 2-kb sequence described above may be a transposable element; both the inserted and the recipient copy of this sequence terminate in an 11-bp inverted repeat. Thus, this 11-bp sequence is found four times, the two right-hand and the two left-hand ones being oriented in the same direction with respect to each other (Fig. 3).

The internal 11-bp inverted repeats are flanked by 8-bp direct repeats. If these are the result of insertion of one copy of the 2-kb element into the other, the 8-bp sequence should be found only once in the internal copy of the transposable element. This is borne out by the sequence. No 8-bp duplication flanks the recipient 2-kb structure. We expect that the 8 bp of *Sh* DNA bordering the inserted DNA at the 3' junction will be repeated at the 5' junction of the 30-kb insert. However, this has not yet been cloned. We conclude, therefore, that the 2-kb sequence is indeed a transposable element and that the more complicated 4-kb structure was created by the insertion of one copy of this element, in inverted orientation, into the other one. We will show below that this transposable element is an example of McClintock's *Ds* and we will refer to the structure described here as a double-*Ds*, or *D(s'D')*s.

Relationship to *Dissociation* elements?

It is conceivable that the large, ~30-kb insertion in *sh-m5933* is in fact *Ds*. We argue here, however, that the small, ~2-kb

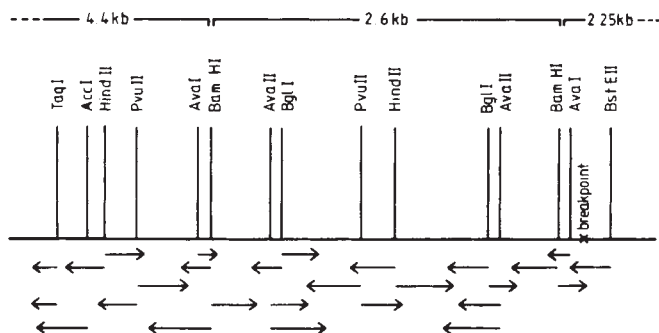


Fig. 1 Sequencing strategy. All sequences were determined (usually twice) by 5' end labelling and chemical degradation according to Maxam and Gilbert³². The arrows indicate the sites of labelling and the extent of sequence read.

structure is *Ds* and that the large insertion in *sh-m5933* consists of several copies of *Ds* and some unique DNA not belonging to this transposable element. Such a conclusion cannot be drawn from the sequence analysis of one element alone; our structure must be compared with other mutations caused by the integration of *Ds*. However, we present the following evidence:

(1) We have isolated a null mutation, *adh1-2F11*, in the gene *Adh1* which encodes alcohol dehydrogenase²². Genetically, this mutation is caused by *Ds*, because it is unstable in the presence of *Ac*, while no reversion to the wild type phenotype is observed in the absence of *Ac*. DNA blotting experiments show the presence of a 1.3-kb insertion in the gene. RNA blotting experiments show an increase in the RNA size corresponding to the insert size within the gene. The enlarged RNA hybridizes to a probe derived from the 4-kb sequence described above, while the wild type RNA does not, indicating that the two inserts in the genes *Sh* and *Adh1* have some DNA sequences in common.

(2) Another mutation, *Adh1-Fm335*, was isolated previously by Osterman and Schwartz²³. By genetic criteria, the mutation is caused by insertion of *Ds*. Peacock and collaborators have sequenced the wild type, mutant and revertant genes^{24,25}, and found that the mutation is caused by the insertion of a 405-bp DNA segment. The 11 nucleotides at the termini are inverted repeats. The sequence found at the termini of the 2-kb transposable element in *sh-m5933* is identical to the inverted repeat first described by Peacock *et al.*²⁴ in the *Adh1* *Ds* insertion. The insertion in *Adh1* is also flanked by a direct repeat of 8 bp. These findings argue that different copies of *Ds* terminate in the same 11-bp repeat and generate 8-bp duplications of the recipient sequence on insertion. Analysis of revertants at the *Adh1* locus²⁴ has demonstrated that the 405-bp insert bounded by 11-bp inverted repeats is a *Ds* element, supporting our contention that the 2-kb structure which we have described is also a *Ds* element, that the inverted 11-bp sequence may be sufficient for transposition by an *Ac*-encoded function of a DNA sequence bracketed by these inverted repeats, and that more complicated structures can be built from the simple unit *Ds*.

(3) Another *Ds*-induced mutation at the *Shrunken* locus, *sh-m6233*, has been analysed by the blotting technique²⁰. A genomic DNA clone containing one terminus of the insertion and adjacent DNA of the *sh* locus has been isolated and the junction sequenced (E. Weck, U. Courage-Tebbe, H.-P.D. and P.S., in preparation). The first 70 nucleotides adjacent to the junction including the 11 nucleotides at the junction, are identical to the sequence of the 2-kb transposable element found in *sh-m5933*.

These biochemical findings are strongly supported by the genetic observation that excision of the 30-kb structure inserted in *sh-m5933* and reversion to the *Shrunken* wild type occurs only in the presence of *Ac*²¹. Thus, the mobile transposable

element-like structure ending in 11-bp termini is probably an example of *Ds*.

Length variation of *Dissociation* elements

From the above, it is clear that *Ds* isolates can differ in length. The *Ds1* of *adh Fm335* sequenced by Peacock *et al.*²⁴ is only 405 bp long, while our unit *Ds* is 2 kb long. The *Ds* in *adh1-2F11* has an intermediate length of ~1.3 kb. Fedoroff *et al.*²⁶ found a *Ds* in a clone of the gene *Waxy* isolated from mutant *wx-m6*, which closely resembles our 2-kb *Ds* by restriction mapping. The finding of *Ds* isolates of different length and the observation that *Ds* is deficient in several functions such as excision from a gene locus, transposition and production of chromosome breaks at the site of insertion, but that it can be complemented *in trans* by *Ac* suggests a hypothesis for the origin of different *Ds* copies⁸. It is conceivable that these arise from *Ac* by internal deletions which both reduce the length and abolish the function. Fedoroff *et al.*²⁶ have isolated *Ac* from another *Waxy* mutant, *wx-m9*, and found that it is ~4.2 kb long. Heteroduplex analysis and restriction mapping has shown that *Ds6* of *wx-m6*, is an internal deletion of *Ac* of *wx-m9*.

Comparison of the *Ds* sequenced by Peacock *et al.*^{24,25} and *Ds5933* shows that the situation may be more complex. Apart from the 11-bp repeats at the termini, there is little sequence similarity. Thus, *Ds* can either be formed from *Ac* by a multi-step event, or there exist different families of *Ac* and *Ds*²⁷.

Evolution of double *Ds*

The creation of independent *Ds* copies by deletions from *Ac* has been discussed above. Another possible means of evolution of different *Ds* structures is evident from the structure of the double *Ds* at the 3' junction in *sh-m5933*. From Fig. 3 it is apparent that different structures bordered by two inverted repeats of 11 bp could be transposed from the double *Ds* (these are indicated by brackets). The data obtained by DNA blotting analysis of the inserted DNA adjacent to the 5' junction in *sh-m5933* are compatible with this assumption²¹. The restriction pattern of the mutant and of several of its revertants is most easily explained if the sequence located at the 5' junction of *Shrunken* and *Ds* DNA is an inverted copy of part of the double *Ds* at the 3' junction (indicated by the bracket marked II in Fig. 3). If this interpretation is correct, the structure at the 5' junction of the insertion consists of 1.5 copies of the 2-kb *Ds*. Part of the 2-kb sequence is present only once at the 5' junction, while the other part is present twice (Fig. 3).

The deletion of part of the structure in one of the revertants, *sh-r5*, and the restriction pattern of the non-deleted part are compatible with the assumption that a unit *Ds* has been excised (Fig. 3). We assume that the remaining half *Ds* element carries the characteristic 11-bp sequence at one terminus only. This will have to be checked by cloning and sequencing. We anticipate that this structure can no longer be excised. However, chromosome breaks at this site are still detected²¹. It is possible that the occurrence of such breaks is not dependent on the presence of the 11-bp sequence at both termini.

Other sequence features of *Ds*

The sequence described here has a few peculiarities, including similarities to other transposable elements, direct or inverted short duplications and insertion sites, repeated oligonucleotides within *Ds* as well as an open reading frame. In none of these cases is a functional significance known.

Termini of *Ds*: Several transposable elements have the dinucleotide TG at the 5' terminus of their long terminal repeat. In some of the retroviruses, the characteristic TG is preceded by the dinucleotide AA²⁸. The hexanucleotide ATGAAA is shared between the termini of the murine leukaemia and sarcoma viruses and *Ds*. In the case of *Ds*, this represents the inner hexanucleotide of the characteristic terminal 11-bp sequence.

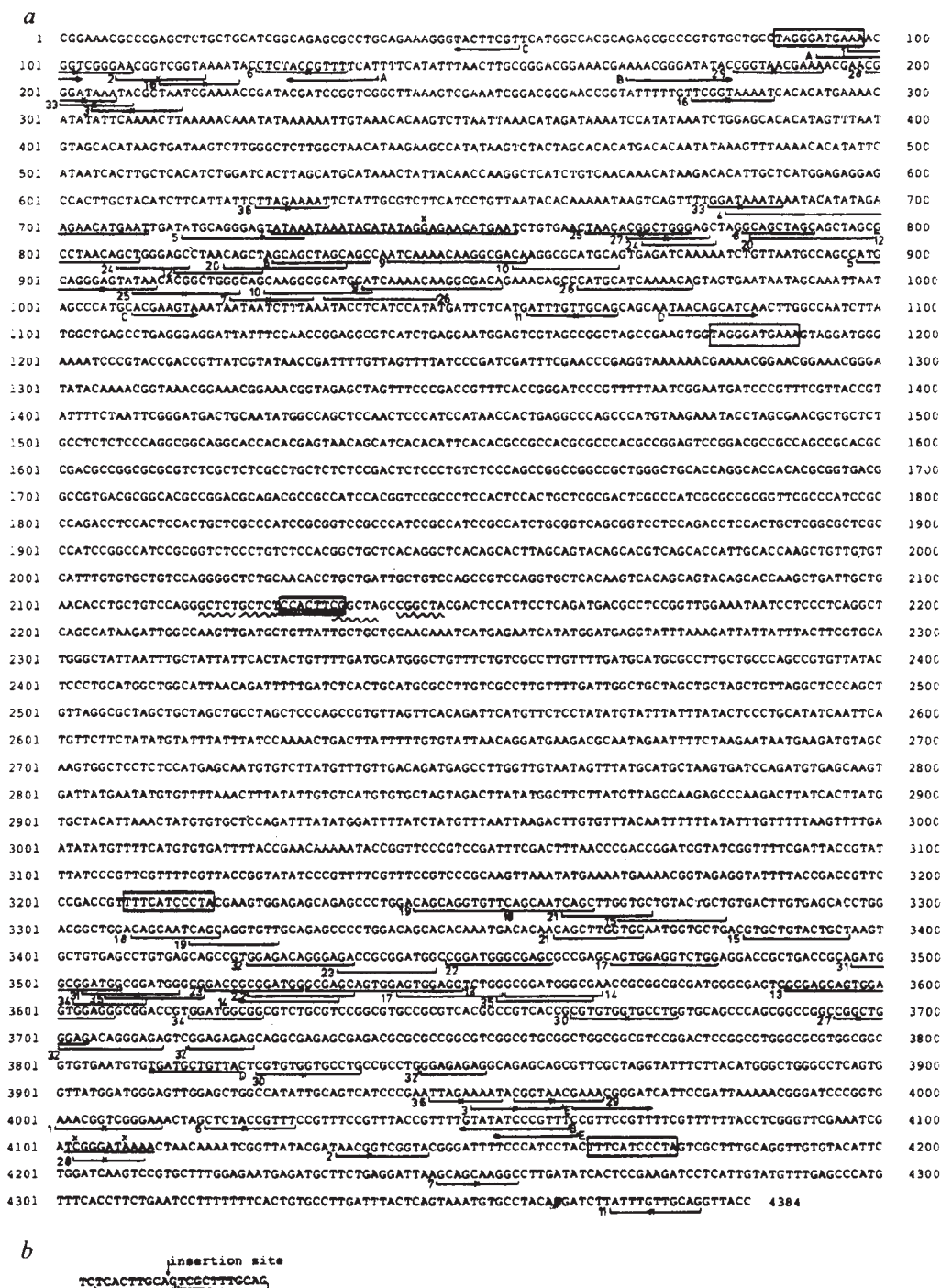


Fig. 2 *a*, Sequence of double *Ds* from *sh-m5933*. The sequence of the transposable element-like structure starts at position 88 and ends at position 4,175. The sequence adjacent to these termini is an 8-bp direct repeat, CGAAGTGG, which is found only once at position 2,128–2,135. The 11-bp inverted repeats and the internal 8-bp insertion site are boxed. The direct repeats (numbered) and the inverted repeats (lettered arrows) are indicated by underlining the outer sequence only. The region between positions 3,960 and 4,080 contains eight copies of CCGTTT, four copies of AAACGG and several near relatives of this sequence. The outer and the inner sequence of *Ds* are identical except for the 8-bp duplication in the outer sequence flanking the insertion of the inner one. At position 955 an ambiguity exists, due to differences in reading the sequencing gels of the opposite strands. Between positions 955 and 956, a C may have to be inserted. The same is true between positions 2,351 and 2,352, where a G might have to be inserted. *b*, Insertion site of *Ds* in *sh-m5933*.

The three adenine residues at the end of the sequence form part of a pentanucleotide AAAGA which is also found at the terminus of the FB transposon in *Drosophila*²⁹. Another similarity between *FB* and *Ds* will be described below. Similarities between the termini of the maize transposon *Cin1* and the *Drosophila* transposon *copia* have been described by Saedler *et al.*³⁰. In this case, the first pentanucleotide was shared between the two transposons.

Many transposable DNA elements terminate in inverted repeats which can be mismatched. We have screened the

sequence inwards from the 11-bp terminal inverted sequences, and have found two penta- and one hexanucleotide on one side which are found in inverted orientation near the other terminus, and which are separated from each other by no more than 10 bp. We have also observed a perfect 11-bp repeat on the left-hand side of *Ds* near one end, at positions 93–103 and 129–139. Near the right-hand end of *Ds* we find a mismatched inverted repeat at positions 4,102–4,112 and 4,165–4,175. These two inverted repeats provide pairing capabilities alternative to that between the two opposite ends of the *Ds* elements.

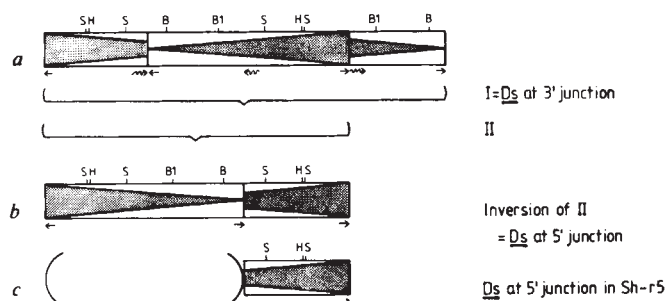


Fig. 3 Schematic representation of double *Ds* and derivatives of it. **a**, Double *Ds* at the 3' junction of *sh-m5933*, as sequenced by us. The solid arrows below the figure indicate the 11-bp inverted repeats (not to scale). The wavy arrows indicate the 8-bp duplication. The brackets indicate two different structures that can presumably be transposed. **b**, Hypothetical structure of *Ds* at the 5' junction of *sh-m5933* both in the complete gene and in the duplication, as derived from restriction mapping on the assumption that structures having 11-bp inverted repeats at their termini can be transposed. **c**, Hypothetical structure of *Ds* at the 5' junction in the duplication of revertant *sh-r5*. This structure is deduced from restriction mapping, which shows a decrease in fragment size for several restriction enzymes cleaving outside *Ds* and a smaller decrease of an *SphI* fragment which is in accord with the map given here. (B = *Bam*HI, S = *Sph*I, B1 = *Bgl*I, H = *Hind*II.)

Insertion sites: In *sh-m5933* there are two *Ds* insertion sites, one located near the centre of *Ds*, where one *Ds* copy is inserted into the other one, the other localized in the sucrose synthase gene, where the 30-kb sequence is inserted. Both of the insertion sites are characterized by short sequence duplications. These are present before the insertion and must not be confused with the 8-bp duplications created by the insertion. In the case of the insertion in the sucrose synthase gene, the duplication is 11 bp long, with one mismatch and one base deletion. The insertion site of *Ds* into *Ds* is characterized by the presence of one hexanucleotide and one pentanucleotide as direct repeats. In both cases, the insertion site is located in one of the duplicated oligonucleotides (see Fig. 2a, wavy lines, and Fig. 2b).

Direct and inverted repeats: The *Ds* sequence is characterized by the presence of many direct and several inverted repeats of variable length. The perfect repeats exceeding 8 bp are indicated in Fig. 2 by underlining. There are additional examples of slightly mismatched repeats which have not yet been analysed in detail.

The position of the direct repeats is non-random. Many are separated by short stretches of DNA, others are adjacent to each other. Some pairs of repeats partially overlap others. We cannot suggest a mechanism for the origin of these repeats apart from the vague notion that in some *Ac-Ds* functions replication and/or recombination events may be involved that are not encountered regularly with other DNA segments.

The repeats are found in most of the *Ds* sequence; only the regions from nucleotides 310–670 and 1,080 to 1,180 are rela-

tively free of them. These regions do not coincide with open reading frames. The 609-bp open reading frame mentioned below is full of repeated sequences. The DNA segments containing the repeated sequences may provide an explanation for the observation that independently isolated *Ds* elements can be of different length and may be derived from *Ac* by internal deletions. Whether a sequence variability generated by recombination events in *Ds* is the basis for the known 'states' of *Ds*^{2,9} may be the subject for future experiments.

Repeated hexanucleotide: Near one end of *Ds*, a region of ~230 bp is characterized by the presence of 15 copies of the hexanucleotide CCGTTT and close relatives of it, which differ from the hexanucleotide by one base substitution. Four of these hexanucleotides are oriented opposite to the others. Some of them lie adjacent to one another, while others are separated by one or a few nucleotides. A similar structure has been described for the *FB* transposon of *Drosophila*²⁹, in which the repeated structure is a decanucleotide. It shares the first five nucleotides, CGTTT, with our hexanucleotide, which may be coincidental. Given the increasing number of similarities between *Drosophila* and maize transposons (see above) the possibility of horizontal transmission of DNA between these two species should be considered.

Open reading frames: The only sizeable open reading frame, starting with an ATG, is located between nucleotides 3,373 and 3,981, possibly encoding a protein of 203 amino acids. The sequence around the ATG at position 3,373, GNNATGG, has been observed to be a translation start site in animals but not yet in plants³¹. We have observed no promoter-like structure preceding it. Whether this open reading frame is transcribed and/or translated is unknown.

Conclusions

The DNA sequence described here resembles other transposons in having inverted DNA repeats at its termini and in creating short direct DNA duplications at the site of insertion. The DNA sequence between the 11-bp inverted repeats does not contain long open reading frames. Whether the abundance of direct and inverted repeats in this structure has a functional significance is unknown. Rearrangements involving these sequences are conceivable and may form the basis for 'changes in state'.

The sequence dissimilarity between *Ds1*²⁵ and *Ds5933* described here suggests that DNA located between the 11-bp inverted repeats is transposed passively. This fact may be exploited in gene transfer experiments. The observation that sequences hybridizing to *Ds5933* are present throughout the maize genome in multiple copies²⁰ will make the identification of particular genes mutated by the insertion of *Ds* difficult, but not impossible. Experiments along these lines are in progress in our laboratory.

We thank K. Stüber for invaluable help in the analysis of DNA sequences and N. Fedoroff and J. Peacock for communicating unpublished results. The work was supported by Deutsche Forschungsgemeinschaft through SFB 74.

Received 4 August; accepted 28 September 1983.

- McClintock, B. *Carnegie Instn. Wash. Yb.* **45**, 176–186 (1946).
- McClintock, B. *Cold Spring Harb. Symp. quant. Biol.* **16**, 13–47 (1951).
- Jordan, E., Saedler, H. & Starlinger, P. *Molec. gen. Genet.* **100**, 296–306 (1968).
- Shapiro, J. A. *J. molec. Biol.* **40**, 93–105 (1969).
- Starlinger, P. & Saedler, H. *Biochimie* **54**, 177–185 (1972).
- Cold Spring Harb. Symp. quant. Biol.* **45**, (1980).
- Mobile Genetic Elements* (ed. Shapiro, J. A.) (Academic, New York, 1983).
- Fincham, J. R. S. & Sastry, G. R. K. *A. Rev. Genet.* **8**, 15–50 (1974).
- McClintock, B. *Brookhaven Symp. Biol.* **18**, 162–184 (1965).
- McClintock, B. *Dev. Biol. Suppl.* **1**, 84–112 (1967).
- Reif, H. J. & Arber, W. *Cold Spring Harb. Symp. quant. Biol.* **45**, 40–45 (1981).
- Saint-Girons, I., Fritz, H. J., Shaw, C., Tillmann, E. & Starlinger, P. *Molec. gen. Genet.* **183**, 45–50 (1981).
- Rubin, G. M. & Spradling, A. C. *Science* **218**, 348–353 (1982).
- Spradling, A. C. & Rubin, G. M. *Science* **218**, 341–348 (1982).
- Döring, H. P., Ehring, R., Geiser, M., Wöstemeyer, J. & Starlinger, P. in *The Plant Genome* (eds Davis, D. R. & Hopwood, D. A.) 73–78 (John Innes, Norwich, 1980).
- Chourey, P. S. & Nelson, O. E. *Biochem. Genet.* **14**, 1041–1055 (1976).
- McClintock, B. *Carnegie Instn. Wash. Yb.* **51**, 212–219 (1952).
- McClintock, B. *Carnegie Instn. Wash. Yb.* **52**, 227–237 (1953).
- Burr, B. & Burr, F. A. *Cell* **29**, 977–986 (1982).
- Geiser, M. et al. *EMBO J.* **1**, 1455–1460 (1982).
- Courage-Tebbe, U., Döring, H. P., Fedoroff, N. & Starlinger, P. *Cell* **34**, 383–393 (1983).
- Döring, H. P. et al. *Molec. gen. Genet.* **193**, 199–204 (1984).
- Osterman, J. C. & Schwartz, D. *Genetics* **99**, 267–273 (1981).
- Peacock, W. J. et al. *Proc. Miami Winter Symp.* (Academic, New York, in the press) (1983).
- Sutton, W. D. et al. *Science* (submitted).
- Fedoroff, N., Wessler, S. & Shure, M. *Cell* **35**, 235–242 (1983).
- Sachs, M. M., Peacock, W. J., Dennis, E. S. & Gerlach, W. L. *Maydica* **XXVIII**, 289–302 (1983).
- Varmus, H. E. in *Mobile Genetic Elements* (ed. Shapiro, J. A.) 411–503 (Academic, New York, 1983).
- Potter, S. S. *Nature* **297**, 201–204 (1982).
- Saedler, H. et al. *5th John Innes Symp.* 107–116 (1982).
- Messing, J. et al. in *Genetic Engineering of Plants* (eds Kusuge, T., Meredith, C. P. & Hollaender, A.) 221–227 (Plenum, New York, 1982).
- Maxam, A. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560–564 (1977).

Mode of proviral activation of a putative mammary oncogene (*int-1*) on mouse chromosome 15

Roel Nusse*, Albert van Ooyen*, David Cox†, Yuen Kai T. Fung‡ & Harold Varmus‡

* Department of Virology, Netherlands Cancer Institute, Antoni van Leeuwenhoekhuis, Plesmanlaan 121, Amsterdam, The Netherlands

† Department of Pediatrics and ‡ Department of Microbiology, University of California, San Francisco, California 94143, USA

*Most mammary carcinomas induced in C3H mice by the mouse mammary tumour virus (MMTV) bear a new proviral insertion within a highly conserved locus on chromosome 15 called *int-1*. A transcriptional unit within this locus is inactive in all tested normal tissues but expressed at low levels in mammary tumours with proviral insertions positioned on either the 5' and 3' sides of the gene. Transcription of the proviruses proceeds away from *int-1*; thus an indirect mechanism appears to activate expression of this putative oncogene.*

THE proviruses of retroviruses can integrate at many sites in host genomes; thus, like other movable genetic elements, they act as insertion mutagens¹. Two classes of proviral insertion mutations have been described: recessive mutations that disrupt and prevent expression of genes²⁻⁶ and presumably dominant mutations that activate gene expression, leading to induction of tumours by retroviruses without their own oncogenes⁷⁻¹³. For example, during the induction of B-cell lymphoma by the avian leukosis virus (ALV), proviral insertion mutations augment transcription of *c-myc*, the cellular progenitor of a retroviral oncogene⁷, either by provision of a retroviral promoter⁷⁻¹⁰ or by an indirect effect on a host promoter¹⁰.

The proposal that retroviruses lacking oncogenes induce tumours by insertion mutation of host genes suggests a strategy for the isolation of novel oncogenes. By molecular cloning of a single new mouse mammary tumour virus (MMTV) provirus from a C3H mouse mammary tumour, we have isolated a region of the mouse genome that contains an MMTV proviral insertion in the majority of C3H mammary tumours¹³. Although this region lacks homology with any of the available retroviral oncogenes, we have previously argued that it harbours an oncogene, called *int-1*, instrumental in mammary tumorigenesis because (1) MMTV proviruses can integrate at many sites in the host genome (implying that insertions near *int-1* confer a selective growth advantage) and (2) because the insertions are accompanied by the appearance of transcriptional products from *int-1*, a silent domain in non-neoplastic mammary tissue.

We demonstrate here that insertion mutations in the *int-1* locus usually stimulate transcriptional activity by an indirect mechanism: MMTV proviruses are located either upstream from transcribed *int-1* sequences in the opposite orientation or downstream in the same orientation. We also assign the *int-1* gene to mouse chromosome 15 and show that *int-1*, like the progenitors of viral oncogenes, has been conserved during evolution.

Locating and orienting MMTV proviruses

We reported previously¹³ that 18 of 26 C3H mouse mammary tumours contain MMTV proviral DNA within a common region of approximately 20 kilobases (kb). The insertions are distributed with roughly equal frequency on both sides of a portion of the locus that is represented in tumour-specific, 2.6-kb polyadenylated RNA. To assess the mechanisms involved in the activation of the *int-1* transcriptional unit, we have now determined the sites of proviral insertion more precisely and ascertained the transcriptional directions of the proviruses and the *int-1* gene. These studies involved physical mapping of the interrupted *int-1* loci with restriction endonucleases, using the

DNA transfer method and a collection of molecular annealing reagents specific for portions of *int-1* and for the 5' and 3' halves of MMTV proviral DNA. The *int-1* probes were derived by plasmid subcloning of unique sequence DNA from recombinant λ bacteriophages containing overlapping inserts of the *int-1* domain from a library of normal BALB/c mouse DNA. Probe C detects *int-1* RNA in mammary tumours; probe D represents a region located leftward of probe C on the *int-1* map as it is conventionally drawn (see Fig. 3). The probes used to orientate proviral DNA [MMTV-*env* and MMTV-*gag* (Fig. 1)] represent portions of MMTV DNA that reside near the 5' and 3' ends of the provirus but exclude sequences from the long terminal repeats (LTRs). In general, our strategy was to use enzymes that cleaved *int-1* at widely separated sites and divided MMTV proviral DNA into two portions distinguishable with MMTV-*env* and MMTV-*gag* probes.

Provirus on the 5' side of *int-1*

Two enzymes, *Bgl*II and *Kpn*I, proved to be particularly useful for the analysis of MMTV proviruses present to the left of the *int-1* transcriptional unit. (We will show below that transcription of *int-1* proceeds from left to right on the physical map; thus, proviruses inserted to the left of the region represented by probe C can be considered to lie 5' to at least a portion of the transcriptional template.) *Bgl*II cleaves the normal C3H mouse DNA at a leftward site (position -21 on the map in Fig. 3) to generate a 14-kb fragment that anneals with probe C. In addition, *Bgl*II cleaves the MMTV provirus at two sites, one of them near the middle of the element (Fig. 1). Proviral DNA in eight tumours was found to disrupt the normal 14-kb *Bgl*II fragment from one of the two *int-1* alleles. As illustrated for six of these tumours in Fig. 1a and b, the novel fragments also annealed with the MMTV-*gag* probe, indicating that the proviruses are positioned either to the right of the probe C homology in the same transcriptional orientation as *int-1* or to the left of the probe C region in the opposite orientation. In all eight tumours, additional mapping, illustrated here for a few tumours with *Kpn*I and probe D, showed that the latter possibility was correct (Fig. 1c, d). *Kpn*I cleaves the *int-1* domain at the most leftward site we have mapped (position -23) and at position -11, to generate a fragment of 12 kb that hybridizes with probe D. The MMTV provirus is cut once by *Kpn*I. Novel *Kpn*I fragments annealing with probe D (Fig. 1c) were also detected by the MMTV-*env* probe (Fig. 1d), demonstrating that the proviruses are inserted upstream from the probe C homology region in orientations opposite to the transcriptional direction of *int-1*.

Table 1 Segregation of mouse *int-1* and mouse chromosome in Chinese hamster–mouse somatic cell hybrids

Chromosome	Marker enzyme	<i>Int-1</i> /chromosome (no. of clones)				Discordance
		+/+	+/-	-/+	-/-	
1	PEP-3	5	4	1	2	42
2	AK-1	15	2	8	4	34
3	—	2	7	1	2	67
4	PGD	13	5	4	8	30
5	—	2	7	1	2	67
6	TPI-1	3	6	7	5	62
7	GPI-1	4	5	2	1	58
8	GR-1	4	5	1	2	50
9	MOD-1	1	8	1	2	75
10	PEP-2	2	7	2	1	75
11	—	0	9	0	3	75
12	ACP-1	5	4	1	2	42
13	—	4	5	0	3	42
14	NP-1	10	6	5	7	39
15	AS-2	17	0	1	10	4
16	SOD-1	3	6	2	0	73
17	GLO-1	2	7	1	2	67
18	PEP-1	7	2	5	8	32
19	GOT-1	12	6	1	11	23
X	HPRT	15	3	12	0	50
Y	—	0	9	0	3	75

Thirty clones of hybrids formed by polyethylene glycol fusion of mouse spleen cells or macrophages with an established Chinese hamster cell line (380-6), deficient for hypoxanthine phosphoribosyl transferase (HPRT), were analysed for mouse *int-1* and enzyme markers, although not all clones were analysed for every enzyme marker. Twelve of the hybrid clones were also analysed by karyotype analysis as described previously²². The symbols for the marker enzymes, their chromosome assignments and the electrophoretic procedures used to separate the Chinese hamster and mouse enzymes have been previously described²². In general, the presence or absence of a given mouse chromosome, as determined by karyotypic analysis, agreed with the presence or absence, respectively, of the enzyme marker for that chromosome. However, in several instances, the enzyme marker for a given chromosome was present in a clone that karyotypically lacked that mouse chromosome, indicating chromosome breakage and/or rearrangement. The presence of mouse chromosomes 3, 5, 11, 13 and Y was based on karyotype analysis alone.

Provirus on the 3' side of *int-1*

Similar analytical strategies were applied to DNA from 10 tumours previously shown to have insertions between positions 0 and -8 on the *int-1* map; an 11th tumour (number 24), previously thought to lack an *int-1* insertion, was also found to have a provirus within this region during the extended tests. Comparative annealings with *int-1* and MMTV probes (ref. 13 and data not shown) revealed that all 11 proviruses were oriented in the same direction, with transcription proceeding from left to right on the *int-1* map.

A particularly interesting part of this analysis, pertinent to the mechanism of expression of *int-1*, is illustrated in Fig. 2. When DNA from tumour 53 was cleaved with *Bam*HI and annealed with probe C, two novel fragments were detected in digests of each sample, indicating that the insert is positioned within the region detected by the probe. Similar results were obtained with DNA from tumour 11 (not shown). Provirus in three other tumours (numbers 55, 12 and 15) were found to be positioned at very similar sites, about 0.6–0.8 kb to the right of the insertion site in tumour 53 (Fig. 2 and unpublished data). The inserts in these tumours lie within the transcriptional unit, and affect the size of resulting transcripts (see below).

We conclude that 19 of 26 C3H mammary tumours carry MMTV proviruses in the *int-1* locus (Fig. 3). Eight proviruses are distributed between positions -11 and -18; each of these is oriented so that transcription of the provirus proceeds from right to left. Eleven proviruses are located within the region

extending from position -8 to 0, and all are oriented so that transcription proceeds from left to right. Mapping with several endonucleases provides no evidence for deletions within either proviral DNA or *int-1*; the several atypical restriction maps of proviral DNA we encountered are best explained by restriction site polymorphisms (data not shown).

int-1 transcripts

We have previously shown that tumours bearing proviruses in the *int-1* region contain polyadenylated transcripts that anneal with probe C, whereas no RNA homologous to probe C is detectable in normal mammary glands¹³. The abundance of the *int-1* transcripts is not great; less than 10 copies per cell have been estimated by three independent methods (unpublished results).

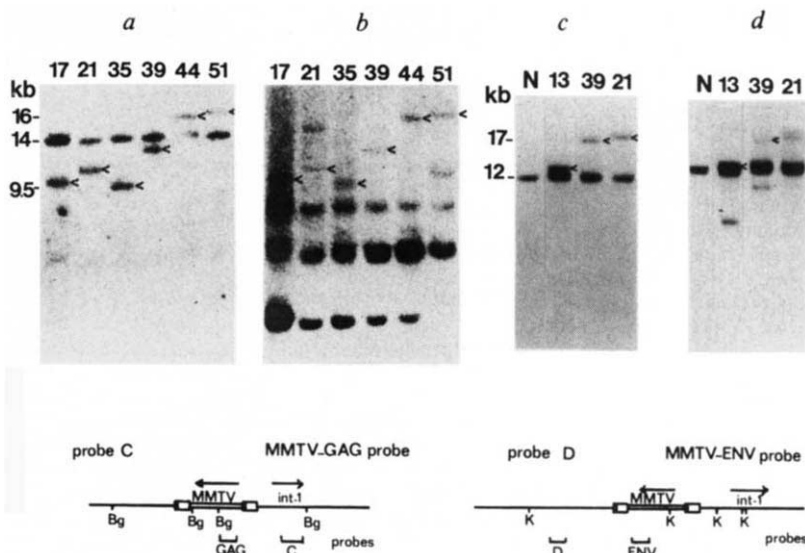
To judge the mechanism by which the insertion mutations activate expression of *int-1*, it was necessary to determine the direction of transcription of the cellular gene. The *Bam*HI-*Eco*RI fragment that constitutes probe C was recloned in both orientations in the bacteriophage vector M13. Restriction mapping of the replicative forms established the orientation of each insert and hence the polarity of strands present in single-stranded virion DNA (data not shown). Radioactive probes were synthesized by random priming on templates provided by two phages, each bearing one of the two strands. Only one of these probes was found to anneal to RNA from tumour 44, previously shown to contain an *int-1* transcript 2.6 kb in length (Fig. 4a); the composition of the probe established the direction of transcription to be from left to right on the *int-1* map as presented in Fig. 3.

The *int-1* transcripts characterized in our earlier report were, like those in tumour 44, approximately 2.6 kb long¹³. In a wider survey of all those tumours from which undegraded polyadenylated RNA could be obtained, we discovered atypical *int-1* transcripts of 3.2 kb in tumour 11 and of 3.8 kb in tumours 12 and 15 (Fig. 4b). In these three tumours, no 2.6-kb *int-1* RNA was found. The conventional 2.6 kb *int-1* RNA species was present in all the other tested tumours with *int-1* insertions (for example, in tumours 24 and 13 in Fig. 4b) and in one tumour (number 14) in which an *int-1* mutation has not been found (unpublished data).

The probable composition of the larger, atypical RNA species was suggested by an intriguing correlation with the site of insertion: in tumours 11, 12 and 15, integration of proviral DNA occurred within or near the right-hand boundary of the region represented by probe C (see Fig. 2). (Suitable RNA samples were not available from the other tumours, 53 and 55, with insertions in this region.) If the proviruses were located within the transcribed domain, upstream from the natural polyadenylation site for *int-1* RNA, the 5' LTR would be appropriately positioned to provide a surrogate polyadenylation site. The enlarged size of the *int-1* transcript could then be ascribed to the length of the U3 region of the MMTV LTR (~1.2 kb)^{14–16}. A similar situation has been encountered in an avian leukosis virus (ALV) insertion mutant affecting the chicken *c-myc* locus¹⁰.

It was not possible to test this explanation by attempting to anneal a labelled probe for the MMTV LTR directly to gel fractionated RNA, because virus-specific RNA transcribed in a normal fashion from proviral DNA was vastly more abundant than transcripts from *int-1*. We attempted to circumvent this problem by using the 'sandwich hybridization' procedure^{10,16} to test for RNA molecules containing MMTV U3 sequences covalently linked to *int-1* sequences. Restriction fragments of phage DNA containing a portion of the *int-1* domain were fractionated in an agarose gel, transferred to nitrocellulose filters, and annealed with unlabelled RNA from tumour 12 (containing a 3.2 kb transcript) or from tumour 13 (with a conventional 2.6 kb transcript). After removal of unannealed RNA, the filter was incubated with a ³²P-labelled *Pst*I fragment containing virtually all of the MMTV LTR. One autoradio-

Fig. 1 MMTV proviruses on the 5' side of the *int-1* transcriptional unit are in the transcriptional orientation opposite to that of *int-1*. DNA samples (10 µg) from mammary tumours (numbers indicated on top) were digested with *Bgl*II (Bg, panels *a* and *b*) or *Kpn*I (K, panels *c* and *d*), electrophoresed in 0.8% agarose gels, and transferred to nitrocellulose¹³. The filters were first incubated with probe C (*a*) or probe D (*c*), washed and exposed to X-ray film. Subsequently, the probes were removed from the filters and the filters were incubated with MMTV-*gag* probe (*b*) or with MMTV-*env* probe (*d*). The origin of the probes is indicated in the diagrams at the bottom of the figure and in Fig. 3; MMTV-*gag* (3.2 kb) is a *Pvu*II-*Bgl*II fragment obtained from a clone of proviral DNA endogenous to GR mice (ref. 56); MMTV-*env* (1.8 kb) is a *Pst*I fragment derived from MMTV (C3H) DNA and cloned into pBR322 (ref. 57). Arrows indicate tumour-specific bands that show rearrangements of the *int-1* locus detected with probes C and D and bands in the same position detected with the MMTV probes.



graphic signal, at the position of the *int-1* fragment homologous to probe C, was observed in the experiment conducted with RNA from tumour 12; no reaction was detected with the sample from tumour 13 (Fig. 4c). From this experiment we conclude that *int-1* and MMTV sequences are covalently joined in transcripts from tumour 12; we presume, but have not proved, that the MMTV sequences are located at the 3' end of the 3.2-kb transcript in this tumour.

Conservation of the *int-1* locus

We have proposed that the *int-1* domain harbours an oncogene that contributes to mammary tumorigenesis following transcriptional activation by proviral integration. However, the locus lacks homology to all tested retroviral oncogenes including *sis*, *ski*, *fos*, *fms* and *yes* (data not shown), as well as the 11 examined previously¹³, and it does not anneal with the NIH/3T3 cell transforming gene detected in mammary tumour DNA by Lane *et al.* (ref. 17 and our unpublished results with M. A. Lane and G. M. Cooper). Moreover, we have been unable to measure *int-1* RNA in a variety of tissues from mature mice (liver, spleen or brain), in cultured mouse cells, or in normal mammary glands from pregnant or lactating females (data not shown). Since no protein product of *int-1* has been identified, we sought support for the idea that the *int-1* region contains a coding domain by testing for conserved *int-1* sequences, homologous to probe C, in the DNA of other vertebrates. As illustrated in Fig. 5, *Bam*HI fragments that anneal effectively with probe C in conditions of only moderately reduced stringency are present in digests of DNA from fish, birds and several mammals, including man. These results are similar to those observed with conserved cellular genes, including the known oncogenes¹⁸. In addition, as found for the progenitors of some retroviral oncogenes¹⁹, sequences homologous to *int-1* can be detected in the genome of *Drosophila melanogaster* (lane 1).

Mouse chromosome 15

Specific karyotypic anomalies, some involving putative oncogenes, have recently been found to be associated with a number of different neoplasms²⁰. Trisomy of mouse chromosome 13, for example, is frequently observed in MMTV-induced mammary tumours²¹. It was therefore of interest to determine the chromosomal location of *int-1*. To assign *int-1* to a specific mouse chromosome, we used probe C to detect restriction fragments containing *int-1* in digests of genomic DNA from 30 Chinese hamster-mouse somatic cell hybrid clones segregating

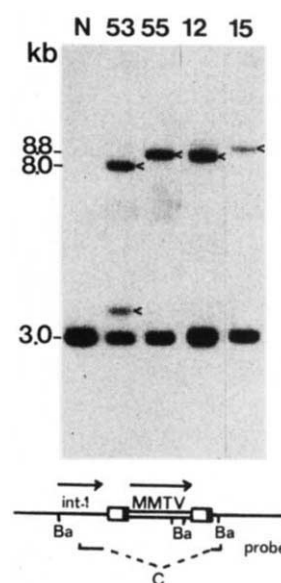
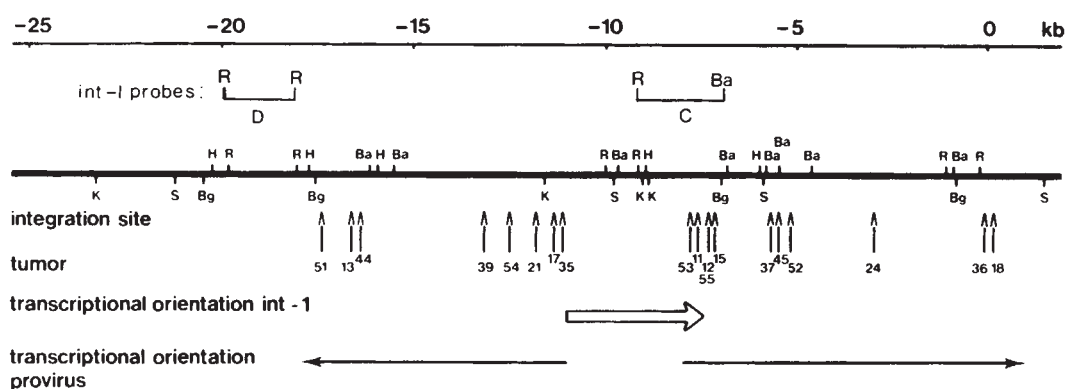


Fig. 2 Proviruses on the 3' side of *int-1* are in the same transcriptional orientation as *int-1*. Tumour DNA (number indicated on top, N: normal DNA) was digested with *Bam*HI (Ba), transferred to nitrocellulose filters and annealed with probe C. The sizes of the arrowed, novel restriction fragments and additional mapping data (not shown) show that the proviruses were integrated as drawn at the bottom of the figure. This illustration shows the provirus from tumour 53, providing an explanation for the two novel *int-1* fragments detected with probe C. Proviruses in the other tumours are located further rightward (see Fig. 3).

mouse chromosomes. These hybrids together contain the full complement of murine chromosomes except for chromosomes 11 and Y (ref. 22). The mouse chromosome complement of each hybrid clone was determined by isozyme and/or karyotype analyses as previously described²². The single *int-1*-specific *Eco*RI fragment of 8.6 kb present in digests of genomic mouse DNA could be easily distinguished from the *int-1*-specific *Eco*RI fragment of 2.8 kb derived from Chinese hamster DNA (not shown). An analysis of the segregation of murine *int-1* and enzyme markers for each of 16 mouse chromosomes (all but 3, 5, 11, 13 and Y) revealed 96% concordant segregation of *int-1* and arylsulphatase A (AS-1), an enzyme marker for mouse chromosome 15, but discordant segregation of *int-1* and the

Fig. 3 A map of the proviral integration sites in the *int-1* locus in C3H mice. Restriction enzyme sites were mapped using C3H and BALB/c chromosomal DNA and DNA from several clones of recombinant bacteriophage containing overlapping inserts from normal BALB/c DNA. Position 0 is defined as the insertion site of the MMTV provirus in tumour 18; this provirus was originally cloned with its flanking DNA to obtain



DNA from the *int-1* locus. Other molecular clones, encompassing positions -22 to +20 on the *int-1* map, were obtained by sequential screening of a Charon 4A bacteriophage library of normal BALB/c mouse DNA with probes from the *int-1* region, as described previously¹³. R, *EcoRI*; Ba, *BamHI*; Bg, *BglII*; K, *KpnI*; S, *SacI*; H, *HindIII*. Probes C and D are fragments of single copy DNA positioned as indicated on the map. Vertical arrows above tumour numbers designate sites of integration. The wide horizontal arrow denotes transcriptional orientation of *int-1* and the approximate extent of the transcribed domain. The narrow horizontal arrows show the direction of transcription of proviruses integrated in the regions demarcated by the arrows.

other enzyme markers (Table 1). When probe C was removed from the DNA blots and the samples reannealed with a probe for the *c-myc* oncogene (data not shown), we observed perfect concordance between *int-1* and the mouse *EcoRI* DNA fragment derived from the *c-myc* locus previously assigned to mouse chromosome 15 (refs 23, 24). This finding, consistent with the isoenzyme results, provided further evidence for the assignment of *int-1* to mouse chromosome 15. Karyotype analyses of 12 selected hybrid clones ruled out an association between *int-1* and chromosomes 3, 5, 11, 13 or Y and were also consistent with the assignment of *int-1* to chromosome 15 (Table 1). The single discordant clone, which was weakly positive for AS-2 but negative for *int-1*, revealed no evidence of a recognizable mouse chromosome 15. Presumably, this hybrid contains a small, non-recognizable fragment of chromosome 15 which includes AS-2 but not *int-1*. Taken together, these data indicate that *int-1* is located on mouse chromosome 15.

Implications

The results described here support our previous contention that tumour induction by MMTV involves insertion mutations that activate a previously unknown gene called *int-1*. At the same time they promote confidence in the general proposition that oncogenic viruses lacking viral oncogenes may initiate tumour growth by insertional mechanisms. This hypothesis fosters an experimental strategy for isolation of oncogenes that is formally analogous to what has been termed 'transposon tagging' in other systems²⁵: identification of the mutagenic provirus in tumour DNA should lead directly to the isolation of a linked, mutated cellular oncogene. These ideas now receive support from studies of avian B-cell lymphomas induced by at least two different retroviruses^{7,11}, erythroleukaemia induced by an avian virus that also causes B-cell lymphoma²⁶, T-cell lymphoma induced by murine leukaemia virus²⁷ and mammary tumours induced by MMTV^{13,28}. At present only circumstantial evidence implicates the affected cellular genes in the oncogenic process: (1) mutation of the same gene in several tumours of the same type; (2) augmented expression and (3) in some cases homology to progenitors of retroviral oncogenes.

Application of the transposon tagging method has led to the discovery of at least two potential mammary oncogenes: the *int-1* gene commonly implicated in tumours arising in C3H mice and the *int-2* domain recently found by Peters *et al.*²⁸ to be interrupted by proviral insertions in about half of the mammary tumours arising in the BR6 mouse strain. These two targets lack evidence of homology with known retroviral oncogenes or with each other, and they are located on different chromosomes (C. Dickson and R. N., unpublished results). In addition, since

substantial numbers of tumours lack insertions in either domain, there may be yet other targets for MMTV insertion mutations.

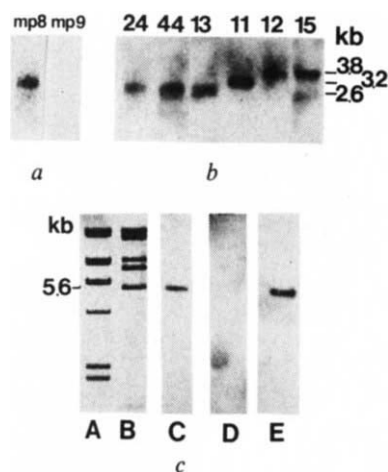
The case for oncogenic roles for *int-1* and *int-2* rests primarily on the argument that MMTV proviral DNA appears capable of integration into many different regions in the mouse genome; hence the repeated finding of proviruses at these two loci implies that the tumours are clonal outgrowths of cells²⁹ with a growth advantage conferred at least in part by the insertion mutations. The case for *int-1* also rests on the evidence that the locus is transcriptionally active in tumours bearing *int-1* mutations but not in normal tissues. Although a protein product of *int-1* has not been identified, a coding function is implied by the conservation of at least part of the transcribed domain during vertebrate evolution (Fig. 5); these findings conform to those obtained with the cellular genes accepted as potential oncogenes by virtue of their homology with retroviral oncogenes.

Our assignment of *int-1* to mouse chromosome 15 does not offer additional insight into its role as an oncogene: karyotypic abnormalities involving this chromosome have not been reported in mouse mammary tumours, and no endogenous MMTV proviruses have been assigned to chromosome 15 (ref. 30). However, an unrelated oncogene, *c-myc*, also present on mouse chromosome 15, recombines with an immunoglobulin locus on chromosome 12 to produce the translocations found in several plasmacytoma lines^{24,31}. Although the distance between *int-1* and *c-myc* on mouse chromosome 15 is not known, mammary tumours with proviral insertions at *int-1* do not display abnormal expression of *c-myc* (unpublished data of authors).

The core of the present report concerns the mechanisms by which MMTV proviruses activate expression of *int-1*. In ALV-induced bursal lymphomas, augmented expression of *c-myc* is most commonly achieved by a mechanism referred to as promoter insertion: ALV proviral DNA on the 5' side of *c-myc* exons in the same transcriptional orientation provides an efficient transcriptional promoter in the form of an LTR⁷⁻¹⁰. However, in the 19 MMTV-induced mammary tumours studied here we have found no instance in which a promoter insertion mechanism activates expression of *int-1*. Instead, MMTV proviruses are positioned either on the 5' side of the transcriptional template for *int-1* in the opposite orientation or on the 3' side in the same orientation. (This rule, however, is not inviolable: in one tumour recently obtained from a BALB/cC3H mouse, a provirus is present on the 5' side of the *int-1* transcriptional domain in the same orientation (R.N., unpublished).)

Previous examples of the two arrangements observed here, in which an LTR cannot serve as promoter for an activated gene, have been described in ALV-induced bursal lymphomas¹⁰. In these few instances, the effect on expression of *c-myc* was ascribed to a retroviral 'enhancer' element, akin to similar elements first identified in papovavirus genomes³²⁻³⁸. Using

Fig. 4 Analysis of *int-1* transcription in mammary tumours. *a*, The direction of transcription. 5 µg of polyadenylated RNA from mammary tumour 44 was subjected to electrophoresis in a 1% agarose gel with 2.2 M formaldehyde, transferred to nitrocellulose filters and hybridized with a strand-specific radioactive probe derived by subcloning probe C (a *Bam*HI-*Eco*RI fragment) in the single-stranded bacteriophage derivatives of M13, mp8



and mp9, in both orientations. The autoradiogram shows the hybridization to 2.6-kb RNA with probes made by transcribing the recombinant M13mp8 DNA with reverse transcriptase in the presence of 32 P-dCTP with oligomers of calf thymus DNA as random primers. A probe made similarly from the recombinant M13mp9 DNA did not detect a transcript. The direction of transcription of the *Eco*RI-*Bam*HI fragment must proceed from the *Eco*RI site to the *Bam*HI site or from left to right on the *int-1* map as shown in Fig. 1. *b*, The length of *int-1* polyadenylated RNA in mammary tumours varies from 2.6 to 3.8 kb. Polyadenylated RNA (5 mg) from mammary tumours with MMTV proviral integrations at the *int-1* locus was subjected to electrophoresis in 1% agarose gels containing 2.2 M formaldehyde¹³ and hybridized with probe C. Tumour numbers are indicated on top (see Fig. 3 for the positions of the proviruses within *int-1*). Tumours 24, 44 and 13 (and also tumours 17, 18, 21, 35, 36, 37, 39, 45) contain a 2.6-kb RNA hybridizing with probe C. Tumour 11 has a 3.2-kb RNA and tumours 12 and 15 have a 3.8-kb RNA species. *c*, Sandwich hybridization demonstrates covalent linkage of *int-1* sequences to MMTV LTR sequences in tumour 12. DNA from a bacteriophage clone covering the normal *int-1* region from positions -20 to -3.4 was digested with *Eco*RI and subjected to electrophoresis (lane B), next to a marker lane (A) containing wild-type λ DNA digested with *Hind*III. DNA was transferred to nitrocellulose filters (several identical filters were prepared from one gel). Direct annealing to radioactive probe C identified a 5.6-kb *Eco*RI fragment (arrow) that contained the *int-1* transcriptional domain (lane C). (This fragment maps from the *Eco*RI site at position -9 in *int-1* to an *Eco*RI linker at position -3.4.) Other filters were first incubated with unlabelled RNA from tumour 13 (lane D) or tumour 12 (lane E). These filters were washed and then hybridized to a radioactive probe from the MMTV LTR (a 1.4-kb *Pst*I fragment containing all but 10 bp of the LTR). Only the filter that was annealed first with the RNA from tumour 12 shows hybridization with the *Eco*RI fragment that also hybridized with probe C. Lanes A and B are photographs of the ethidium-bromide stained gel, the other lanes are autoradiograms.

Methods: 20 µg of poly(A)⁺ RNA was resuspended in 2 ml of an annealing mix consisting of 50% formamide, 0.5 M NaCl, 20 mM PIPES pH 6.8, 5 mM Na₂EDTA, 0.4% SDS, 250 µg ml⁻¹ poly(A), 2× Denhardt's solution, and 200 µg ml⁻¹ carrier yeast RNA. This solution was applied to the filter-bound DNA and incubated for 24 h at 41 °C. The filter was washed twice at 53 °C for 15 min with 0.1×SSC, 0.1% SDS. The filter was blotted dry, annealed to MMTV-LTR probe (2×10⁶ c.p.m.), washed twice for 1 h at 37 °C with 0.1×SSC, 0.1% SDS, dried and autoradiographed.

DNA transformation techniques, retroviral³⁹⁻⁴² and cellular⁴³ sequences with enhancer activity have been shown to act in *cis*, over variable distances and independent of polarity, to improve the efficiency of heterologous promoters.

Our evidence that MMTV DNA can activate the expression of a previously silent gene in tumours has several of the characteristics of enhancement: the mechanism of activation operates in a manner independent of orientation, over distances of at least 10 kb, and in *cis* rather than *trans*. The basis for claiming *cis* activation is our observation that insertions between positions -8 and -6 on the *int-1* map are associated only with

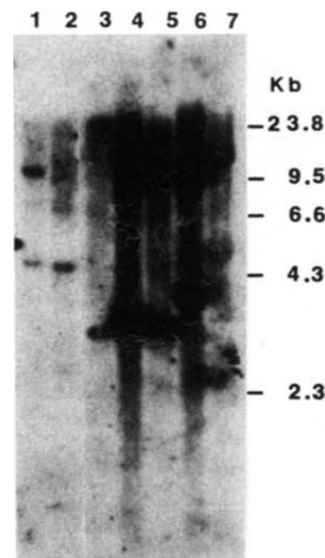


Fig. 5 Sequences homologous to *int-1* are present in all vertebrates and *Drosophila*. Cellular DNA (10–15 µg) from the indicated sources was digested with *Bam*HI, subjected to electrophoresis in an 0.8% agarose gel, transferred to nitrocellulose filters, and annealed with probe C in 37.5% formamide, 3×SSC, at 42 °C. After washing in 2×SSC at 50 °C, the filter was exposed to X-ray film for 6 days. Lane 1, adult *Drosophila melanogaster* (Oregon-S strain); lane 2, fish (*X. helleri*); lane 3, chicken erythrocytes (strain 15×7); lane 4, mouse liver (BALB/c strain); lane 5, rat hepatoma cells (H4IIE line, ATCC); lane 6, domestic cat liver; lane 7, human mammary epithelial cells (HBL100 line; ATCC). Numbers at the right (in kb) refer to the positions of *Hind*III fragments of λ phage DNA present in the same gel.

transcripts of atypically large size (Fig. 4b), despite the persistence of the chromosome bearing a normal *int-1* allele; we conclude that only the physically rearranged locus has been activated. Other aspects of the situation may also be germane to a consideration of mechanism. (1) The observed effect is conversion of an ostensibly silent gene to one expressed at low levels (less than 10 RNA molecules per cell), rather than to amplify expression of a gene normally transcribed. (2) Enhancer functions are likely to be cell specific⁴⁴⁻⁴⁹ and perhaps target gene specific, so it is plausible that the MMTV enhancer works preferentially in mammary cells and upon *int-1*. (3) Initiation of MMTV transcription is greatly stimulated by glucocorticoid hormones⁵⁰, and the hormonal response displays some of the properties of enhancement⁵¹⁻⁵³; hence it is possible that activation of *int-1* is also affected by physiological doses of hormone. This final point should be amenable to direct test by growing mouse mammary tumour cells with *int-1* insertions in controlled conditions in culture.

The use of an MMTV enhancer function to activate expression of *int-1* could explain the uniform orientations of proviruses on each side of the proposed transcriptional unit (see Fig. 3). Enhancers are thought to act only or preferentially upon proximal promoters^{54,55}; if the MMTV enhancer is located on the 5' side of the initiation site for viral RNA, within the 1.2-kb U3 sequence of the LTR, then proviruses would require the orientations observed to avoid interposing an MMTV promoter between the enhancer and its *int-1* target.

We thank P. Borst, C. Dickson, J. Majors and W. Hayward for discussions and comments on the manuscript, M. Simon, M. Schwab and C. Hammond for materials, M. Paape for technical assistance, N. van Nuland and J. Marinos for preparing the manuscript, and E. Stavnezer, C. Sherr, I. Verma, R. Gallo and M. Yoshida for provision of cloned viral oncogenes. A.v.O. was supported by the Netherlands Cancer Society Koningin Wilhelmina Fonds. This work was supported by grants from the NIH and the American Cancer Society (to H.E.V. and D.C.) and Koningin Wilhelmina Fonds (to R.N.).

Received 22 July; accepted 30 September 1983.

1. Varmus, H. E. in *Mobile Genetic Elements* (ed. Shapiro, J. A.) 411–503 (Academic, New York, 1983).
2. Varmus, H. E., Quintrell, N. & Ortiz, S. *Cell* **25**, 23–36 (1981).
3. Jenkins, N. A., Copeland, M. G., Taylor, B. A. & Lee, B. K. *Nature* **293**, 370–374 (1981).
4. Copeland, N. G., Jenkins, N. A. & Lee, B. K. *Proc. natn. Acad. Sci. U.S.A.* **80**, 247–249 (1983).
5. Jaenisch, R. *et al.* *Cell* **32**, 209–216 (1983).
6. Kuff, E. L. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **80**, 1992–1996 (1983).
7. Hayward, W. S., Neel, B. G. & Astrin, S. M. *Nature* **290**, 475–480 (1981).
8. Neel, B. G., Hayward, W. S., Robinson, H. L., Fang, J. & Astrin, S. M. *Cell* **23**, 323–324 (1981).
9. Payne, G. S. *et al.* *Cell* **23**, 311–322 (1981).
10. Payne, G. S., Bishop, J. M. & Varmus, H. E. *Nature* **295**, 209–213 (1982).
11. Noori-Daloii, M. R., Swift, R. A., Kung, H. J., Crittenden, L. B. & Witter, R. L. *Nature* **294**, 574–575 (1981).
12. Fung, Y. K. T., Fadly, A. M., Crittenden, L. B. & Kung, K. J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3418–3422 (1981).
13. Nusse, R. & Varmus, H. E. *Cell* **31**, 99–109 (1982).
14. Shank, P. R., Cohen, J. C., Varmus, H. E., Yamamoto, K. R. & Ringold, G. M. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2112–2116 (1978).
15. Donchower, L. A., Huang, A. L. & Hager, G. L. *J. Virol.* **37**, 226–238 (1981).
16. Dunn, A. R. & Hassel, J. A. *Cell* **12**, 23–36 (1977).
17. Lane, M. A., Sauten, A. & Cooper, G. M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5185–5189 (1981).
18. Bishop, J. M. & Varmus, H. E. in *The Molecular Biology of Tumor Viruses. Pt III* (eds Weiss, R. A., Teich, N., Varmus, H. E. & Coffin, J. M.) 999–1109 (Cold Spring Harbor Laboratory, New York, 1982).
19. Shilo, B. Z. & Weinberg, R. A. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6789–6797 (1981).
20. Rowley, J. D. *Nature* **301**, 290–291 (1982).
21. Dofuku, R., Utakoji, T. & Matsuzawa, A. *J. natn. Cancer Inst.* **63**, 651–656 (1979).
22. Cox, D. R., Sawicki, J. A., Yee, D., Appella, E. & Epstein, C. J. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1930–1934 (1982).
23. Crews, S., Barth, R., Hood, L., L., Prehn, J. & Calame, K. *Science* **218**, 1319–1321 (1982).
24. Sakaguchi, A. Y., Lalley, P. A. & Naylor, S. L. *Somat. Cell Genet.* **9**, 391–406 (1983).
25. Bingham, P. M., Levis, R. & Rubin, G. M. *Cell* **25**, 693–704 (1981).
26. Fung, Y. K., Lewis, W. G., Crittenden, L. B. & Kung, H.-J. *Cell* **33**, 357–368 (1983).
27. Tsihiis, P. N., Gunter-Strauss, P. & Hu, L. F. *Nature* **302**, 445–449 (1983).
28. Peters, G., Brooke, S., Smith, R. & Dickson, C. *Cell* **33**, 369–377 (1983).
29. Cohen, J. C., Shank, P., Morris, V. L., Cardiff, R. & Varmus, H. E. *Cell* **16**, 333–345 (1978).
30. Coffin, J. M. in *The Molecular Biology of Tumor Viruses. Pt III* (eds Weiss, R. A., Teich, N., Varmus, H. E. & Coffin, J. M.) 1109–1203 (Cold Spring Harbor Laboratory, New York, 1982).
31. Sheng Ong, G. L., Keath, B. J., Piccoli, S. P. & Cole, M. D. *Cell* **31**, 443–452 (1982).
32. Banerji, J., Rusconi, S. & Schaffner, W. *Cell* **27**, 299–308 (1981).
33. Moreau, P., Hen, R., Everett, R. & Gaub, M. P. *Nucleic Acids Res.* **9**, 6047–6069 (1981).
34. DeVilliers, J. & Schaffner, W. *Nucleic Acids Res.* **9**, 6251–6264 (1981).
35. Capecchi, M. R. *Cell* **22**, 479–488 (1980).
36. Weiher, H., Konig, M. & Gruss, P. *Science* **219**, 626–631 (1983).
37. Lusky, M., Berg, L., Weiher, H. & Botchan, M. *Molec. Cell Biol.* **3**, 1108–1122 (1983).
38. Fromm, M. & Berg, P. *Molec. cell Biol.* **3**, 991–999 (1983).
39. Levinson, B., Khoury, G., Vande Woude, G. & Gruss, P. *Nature* **295**, 568–572 (1982).
40. Kriegler, M. & Botchan, M. *Molec. cell Biol.* **3**, 325–339 (1983).
41. Jolly, P. J., Esty, A. C., Sbramani, S., Friedmann, T. & Verma, I. M. *Nucleic Acids Res.* **11**, 1855–1872 (1983).
42. Luciw, P., Bishop, J. M., Varmus, H. E. & Capecchi, M. *Cell* **33**, 705–716 (1983).
43. Conrad, S. E. & Botchan, M. *Molec. cell Biol.* **2**, 949–965 (1982).
44. Fujimura, F. K., Deininger, P. L., Friedman, T. & Linney, E. *Cell* **23**, 809–814 (1981).
45. Katinka, M., Vasseur, M., Montreau, N., Yaniv, M. & Blangy, D. *Nature* **290**, 720–722 (1981).
46. Sekikawa, K. & Levine, A. J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1100–1104 (1981).
47. Laimins, L. A., Khoury, G., Gorman, C., Howard, B. & Gruss, P. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6453–6456 (1982).
48. Gillies, S. D., Morrison, S. L., Oi, V. T. & Tonegawa, S. *Cell* **33**, 717–728 (1983).
49. Banerji, J., Olson, L. & Schaffner, W. *Cell* **33**, 729–740 (1983).
50. Varmus, H. E., Ringold, G. & Yamamoto, K. R. in *Glucocorticoid Hormone Action* (eds Baxter, J. D. & Rousseau, G. G.) 253–289 (Springer, New York, 1979).
51. Chandler, V. L., Maler, B. A. & Yamamoto, K. R. *Cell* **33**, 489–499 (1983).
52. Hynes, N., Van Ooyen, A. J. J., Kennedy, N., Herrlich, P., Ponta, H. & Groner, B. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3637–3641 (1983).
53. Majors, J. E. & Varmus, H. E. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5866–5870 (1983).
54. Waslylyk, B., Waslylyk, C., Augereau, P. & Chambon, P. *Cell* **32**, 503–514 (1983).
55. DeVilliers, J., Olson, L., Banerji, J. & Schaffner, W. *Cold Spring Harbor Symp. Quant. Biol.* **47**, 911–920 (1982).
56. Hynes, N. E., Kennedy, N., Rahmsdorf, V. & Groner, B. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2038–2041 (1981).
57. Majors, J. E. & Varmus, H. E. *Nature* **289**, 253–258 (1981).

Correlation between segmental flexibility and effector function of antibodies

V. T. Oi, T. M. Vuong, R. Hardy, J. Reidler, J. Dangl, L. A. Herzenberg & L. Stryer

Department of Structural Biology, Sherman Fairchild Center and the Department of Genetics, Stanford University School of Medicine, Stanford, California 94305, USA

Mouse monoclonal anti-dansyl antibodies with the same antigen-binding sites but different heavy chain constant regions were generated. The extent of segmental flexibility in times of nanoseconds and the capacity to fix complement were greatest for IgG2b, intermediate for IgG2a, and least for IgG1 and IgE. Hence, the effector functions of immunoglobulin isotypes may be controlled in part by the freedom of movement of their Fab arms.

IMMUNOGLOBULIN G is a flexible Y-shaped molecule. The two antigen-binding Fab units of IgG are joined to an Fc unit at a hinge that allows the angle between the Fab parts to vary over a broad angular range, as shown by hydrodynamic¹, electron microscopic^{2,3}, and X-ray crystallographic studies^{4–6}. Nanosecond fluorescence polarization measurements have demonstrated that immunoglobulin molecules exhibit segmental flexibility in the nanosecond time range^{7–9}. Segmental flexibility is likely to be important in enabling immunoglobulins to bind optimally to multivalent antigens and to carry out certain effector functions. Studies of polyclonal antibody populations have suggested that immunoglobulin classes differ in their degree of segmental flexibility^{10–11}. We have explored the relationship between hinge motion and effector function in specially constructed families of homogeneous immunoglobulins. Mouse monoclonal anti-dansyl antibodies with the same antigen-combining sites but different heavy chain constant regions (Fig. 1) were generated by selecting somatic variants in hybridoma cell lines. The extent of segmental flexibility in times of nanoseconds (measured by fluorescence spectroscopy) and the capacity to fix complement were greatest for IgG2b, intermediate for IgG2a, and least for IgG1 and IgE. Hence, the effector functions of immunoglobulin isotypes may be controlled in part by the freedom of movement of their Fab arms.

In most antibody-producing hybridoma cell lines, variant cells arise which produce antibody containing a different heavy chain¹². For example, cells producing IgG2b sometimes arise from a cell line producing IgG1. The frequency of these heavy-chain switch variants is ordinarily low (10^{-5} – 10^{-6} per generation). These rare cells can be separated by fluorescence-activated cell sorting on the basis of the appearance of a different heavy chain on their cell surface^{13,14}. In these switch variants, the light chain remains the same, whereas V_H becomes joined to a different C_H . Consequently, the antigen-combining site is the same as in the parent, despite the change in the heavy-chain constant region. Many of these switch variants are stable and can be cloned.

We chose to generate a family of homogeneous mouse anti-dansyl (DNS) antibodies for two reasons. First, the fluorescence emission spectrum of the bound dansyl chromophore is very responsive to the polarity of its environment¹⁵, making it a convenient and sensitive indicator of whether the antigen-combining site of the switch variant antibody is in fact the same as that of the parent line. Second, the bound dansyl chromophore has an excited state lifetime suitable for determining the segmental flexibility of immunoglobulin isotypes by nanosecond fluorescence polarization spectroscopy¹⁶. Variant hybridoma cell lines producing different immunoglobulin isotypes were

Fig. 1 Schematic diagram of immunoglobulin G. In each family of isotypes, the V_L , C_L and V_H regions are the same. The hinge is located between the C_H1 and C_H2 domains.

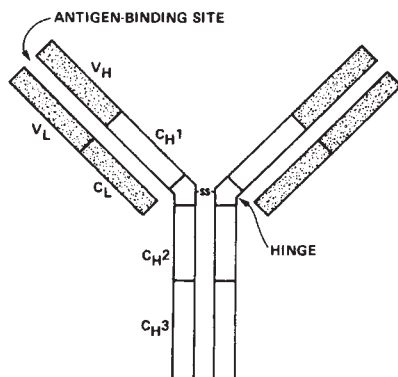


Table 1 Mean rotational correlation times of immunoglobulin isotypes and their Fab fragments

	$\langle\phi\rangle$ (ns)	
	No DTT	With DTT
DNS1 family		
Intact immunoglobulin		
IgE	124	88
IgG1	81	58
IgG2a	60	51
IgG2b	47	42
Fab fragment		
IgG1	28	28
IgG2a	28	28
IgG2b	29	29
DNS3 family		
Intact immunoglobulin		
IgG1	93	55
IgG2b	46	45

selected by using affinity-purified goat and rabbit anti-mouse heavy-chain sera and a panel of hybridoma anti-mouse immunoglobulin allotype antibodies¹⁴. IgG2b, IgG2a and IgE anti-DNS antibody-producing cell lines were isolated starting with the mouse IgG1 anti-DNS hybridoma cell line DNS1 (Igh^a allotype, also designated as 27-44). Our previous nanosecond fluorescence study¹⁶ showed that IgG1 from this mouse cell line is more rigid than is polyclonal rabbit anti-dansyl antibody. The cell lineage of the DNS1 family is: IgG1 (27-44) → IgG2b (27-35) → IgGa (27-13) → IgE (27-74). A cell line containing a different IgG1 anti-dansyl combining site, DNS3 (Igh^b allotype, also designated as 44-10), served as the source of another anti-dansyl IgG2b (44-32).

Two-dimensional gel electrophoresis of antibodies^{17,18} produced by the DNS1 family shows that the heavy chain synthe-

sized by each member has the electrophoretic mobility expected for $\gamma 1$, $\gamma 2b$, $\gamma 2a$ and ϵ chains (Fig. 2). Also, these gels show that the light chains produced by this family are identical. Corresponding results were obtained for the DNS3 family. Heavy chains synthesized by these variant cells in the presence of tunicamycin, a known inhibitor of asparagine-linked glycosylation¹⁹⁻²¹, have a greater electrophoretic mobility than do those synthesized in its absence. This finding suggests that each variant isotype is normally glycosylated in the absence of tunicamycin. The fluorescence emission spectra of ϵ -dansyl-L-lysine bound to switch variants of the DNS1 cell line are identical to that of the parent cell antibody (Fig. 3), strongly suggesting that the DNS1 family has the same antigen-combining site. Likewise, both DNS3 isotypes have identical dansyl fluorescence emission spectra. Also, the dissociation constant of the antibody-hapten complex is the same within the DNS family ($K_d = 17 \pm 1$ nM) and the DNS3 family ($K_d = 3 \pm 1$ nM).

Segmental flexibility

The segmental flexibility of these anti-dansyl immunoglobulins was determined by measuring their fluorescence polarization kinetics in the nanosecond time range^{7,16,22}. A solution of immunoglobulin containing bound ϵ -dansyl-L-lysine is excited by a picosecond pulse of vertically-polarized light. Photoselection produces an ensemble of excited chromophores preferentially aligned in the vertical direction. The orientations of these excited molecules then become randomized by rotational brownian motion. The rate of randomization depends on the size and shape of the immunoglobulin, and also on its modes of flexibility. These motions can be monitored by measuring the intensities of the vertically-polarized [$y(t)$] and horizontally-polarized [$x(t)$] components of the fluorescence emission as a function of time following the excitation pulse. The most informative parameter is the emission anisotropy $A(t)$, which is defined as

$$A(t) = \frac{y(t) - x(t)}{y(t) + 2x(t)}$$

The emission anisotropy kinetics of a chromophore rotating in common with a rigid sphere is given by $A(t) = A_0 \exp(-t/\phi)$, where A_0 is the initial value of the emission anisotropy and ϕ is the rotational correlation time. For a nonspherical particle, $A(t)$ is more complex, consisting of up to five exponential terms. For a rigid ellipsoid, $A(t)$ is a sum of three exponential decays that depend on the volume and axial ratio of the ellipsoid and on the directions of the transition moment of the emitting chromophore. A rigorous theory for the emission anisotropy kinetics of rigid and segmentally flexible Y-shaped molecules has not yet been developed. Emission anisotropy data for immunoglobulins have instead been interpreted phenomenologically by comparing observed $A(t)$ curves with those

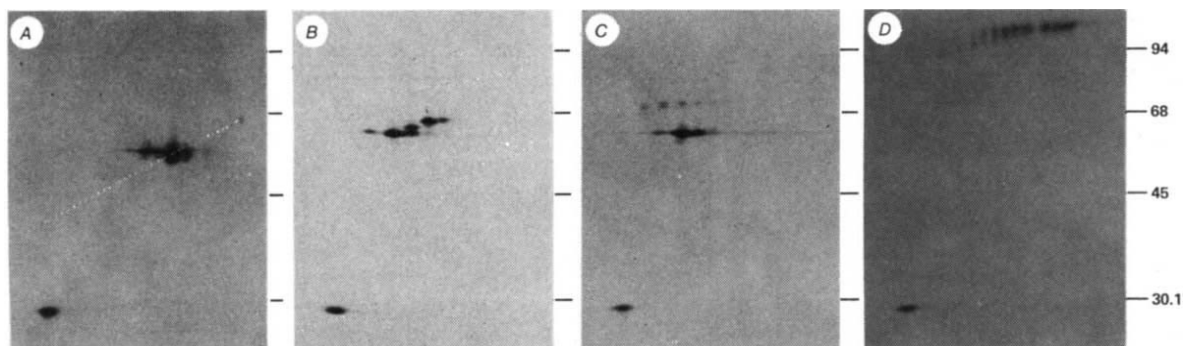


Fig. 2 Two-dimensional gel electrophoresis autoradiogram patterns of antibodies produced by the DNS1 family. A, IgG1; B, IgG2b; C, IgG2a; and D, IgE. The horizontal direction corresponds to a non-equilibrium pH gradient (low pI at the left) and the vertical direction corresponds to sodium dodecyl sulphate polyacrylamide gel electrophoresis^{17,18}. Antibodies were labelled biosynthetically with ³⁵S-methionine. Secreted anti-dansyl antibodies in the culture supernatants were precipitated with monoclonal antibodies specific for $\gamma 1$ (A), $\gamma 2b$ (B), $\gamma 2a$ (C) and a rabbit anti-mouse IgE (D). The spot in the lower left-hand corner appears in the same position in each autoradiogram, providing strong confirmation that the κ light chain is the same in each. The multiple spots in the upper half of each gel are typical for myeloma and hybridoma heavy chains. The heterogeneity probably arises from differences in deamidation and glycosylation.

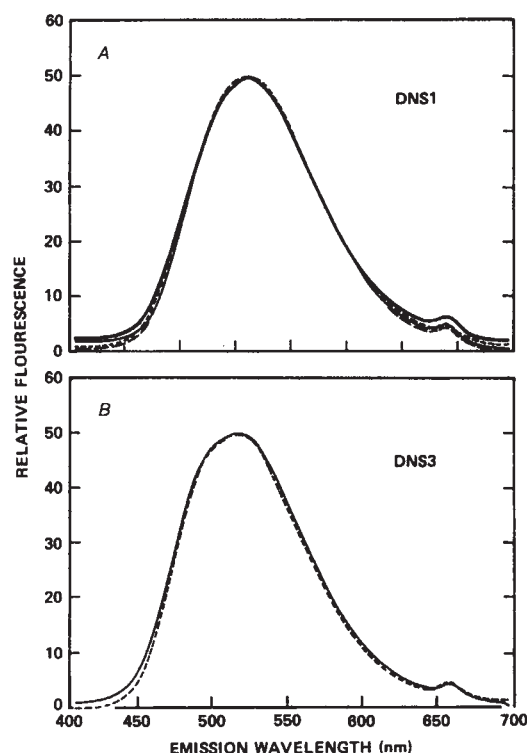


Fig. 3 Fluorescence emission spectra of ϵ -dansyl-L-lysine bound to isotypes of the DNS1 (A) and DNS3 (B) families. The spectra of isotypes within a family are virtually superposable. The excitation wavelength was 340 nm.

calculated for rigid spheres and ellipsoids. In this regard, the mean rotational correlation time $\langle\phi\rangle$ can be informative. $A(t)$ is fitted to a sum of two exponential decays, $a_1 \exp(-t/\phi_1) + a_2 \exp(-t/\phi_2)$, to give $\langle\phi\rangle$, which is equal to $(a_1\phi_1 + a_2\phi_2)/(a_1 + a_2)$. Rigid nonspherical particles have a $\langle\phi\rangle$ greater than ϕ for a rigid sphere of the same volume. Consequently, a particle, irrespective of shape, must be flexible if its $\langle\phi\rangle$ is less than ϕ_{sphere} .

Emission anisotropy kinetics were measured using a mode-locked argon-ion laser and synchronously pumped rhodamine 6G dye laser as the excitation source and a single-photon counting detection system, as described previously¹⁶. Picosecond pulses at 325 nm were produced by doubling the frequency of the 650 nm output of the dye laser. All fluorescence studies were carried out at 20 °C. $A(t)$ curves for the DNS1 family are shown in Fig. 4A. The striking finding is that the IgG1, IgG2a and IgG2b isotypes of this family display markedly different $A(t)$ curves. Their $\langle\phi\rangle$ values are 81, 60 and 47 ns, respectively. These immunoglobulins have virtually identical molecular weights and sedimentation coefficients (6.8, 6.67, and 6.70, respectively; V. Schumaker and M. Phillips, personal communication), which means that their molecular volumes and time-averaged shapes are nearly the same. Furthermore, the Fab fragments obtained from these isotypes have nearly the same $\langle\phi\rangle$ values (Table 1). The orientation of the bound dansyl chromophore with respect to the long axis of the Fab unit is identical or nearly so in these isotypes because they possess identical antigen-combining sites. Thus, the large difference in the $A(t)$ curves (Fig. 4A) are almost certainly due to different extents of segmental flexibility in the intact immunoglobulins. IgG1 is the most rigid, IgG2a is intermediate, and IgG2b is the most flexible of the three isotypes.

IgE is like IgG1 in being more rigid than IgG2a and IgG2b (Fig. 4A). The $\langle\phi\rangle$ value for IgE can be compared with those of the three IgG isotypes by scaling its $\langle\phi\rangle$ to take into account its 190,000 molecular weight²³ compared with 150,000 for IgG.

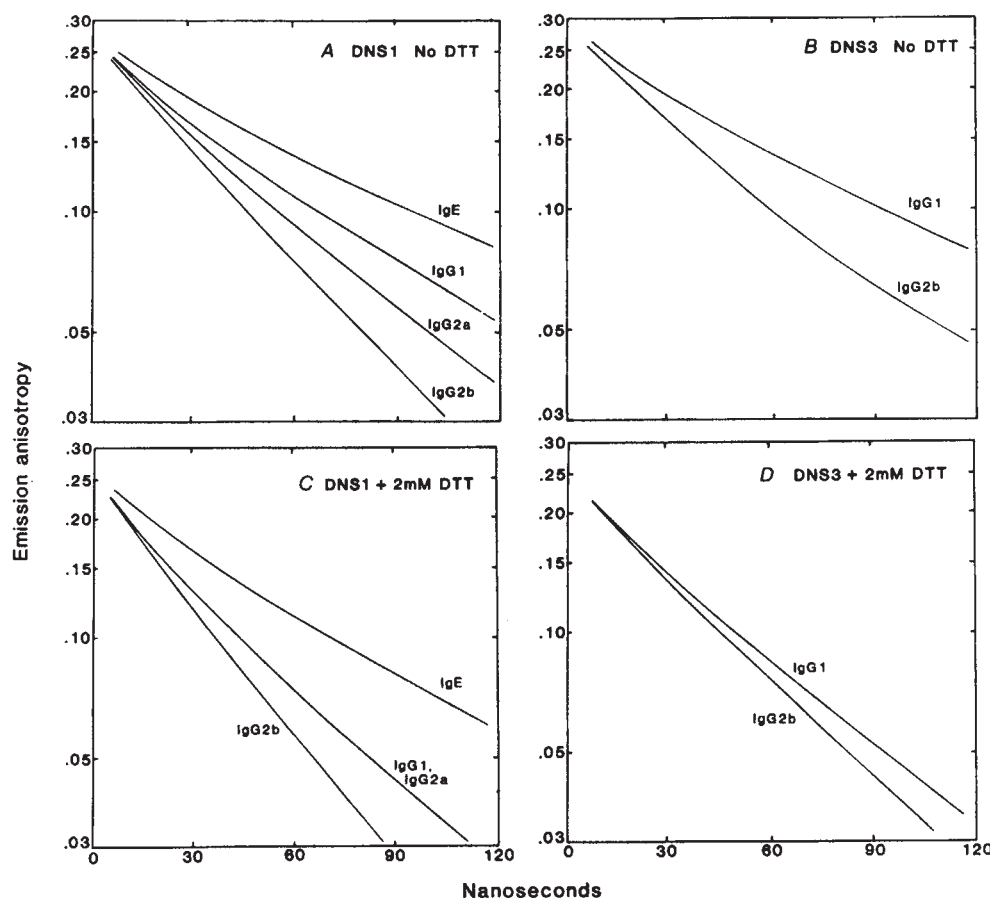


Fig. 4 Nanosecond emission anisotropy kinetics of ϵ -dansyl-L-lysine bound to immunoglobulins of the DNS1 family in the absence (A) and presence (B) of 2 mM DTT; and the DNS3 family in the absence (C) and presence (D) of 2 mM DTT. The temperature was 20 °C.

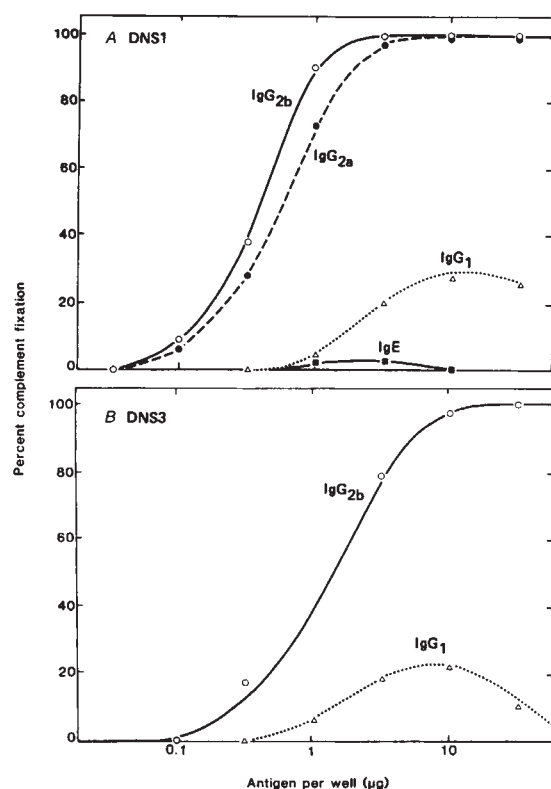


Fig. 5 Complement-fixation by the DNS1 (A) and DNS3 (B) family. A microcomplement consumption assay using ^{51}Cr release from haemolysin-treated sheep erythrocytes similar to that described previously²⁵ was used to measure complement fixation by the DNS1 and DNS3 family. Immunoglobulin (10 μg per 25 μl), DNS-bovine serum albumin dilutions (in 25 μl) and complement (2 CH-50 units in 25 μl) were combined in wells of a 96 well microtitre plate and incubated for 45 min at 37 °C. Next, 50 μl of a 2% solution of activated, ^{51}Cr -loaded erythrocytes was added to each well and after a further 45 min incubation at 37 °C the non-lysed cells were pelleted by centrifugation of the plates. 50- μl aliquots of supernatant were collected from each well and counted on a gamma counter set for ^{51}Cr .

Complement-fixation effectiveness

The ability of these antibody molecules to fix complement was measured using a micro-complement consumption assay²⁵. Dansylated bovine serum albumin served as the antigen. In the DNS1 family, IgG2b is most effective in fixing complement, and IgG2a is also highly effective (Fig. 5A). In contrast, IgG1 fixes less complement and required ten-fold more antigen to trigger complement fixation. IgE exhibits negligible complement-fixation activity. Likewise, in the DNS3 family, IgG2b fixes more complement than does IgG1 and does so at a lower concentration of antigen (Fig. 5B). The isotypes that fix complement most effectively, IgG2a and IgG2b, also exhibit the highest degree of segmental flexibility.

Two families of immunoglobulin isotypes with identical antigen-binding sites were constructed using somatic cell genetics techniques. Nanosecond fluorescence polarization studies of these monoclonal anti-dansyl isotypes have revealed striking differences in their extents of segmental flexibility. IgG2b is the most flexible, IgG2a is less flexible, and IgG1 and IgE are relatively rigid. The degree of molecular motion at the hinge in times of nanoseconds is correlated with the ability of these molecules to fix complement. This correlation may be more than fortuitous. Segmental flexibility may promote complement fixation by enhancing multivalent binding, facilitating the transmission of conformational changes from the Fab units to the effector domain on the Fc unit, or enabling the Fc unit to assume a conformation favourable for complement fixation. It will be very interesting to further explore the relationship between flexibility and effector function by engineering a series of immunoglobulins in which the hinge is deleted, elongated, or modified to alter the rotational mobility of the Fab arms. For example, it would be instructive to study a chimaeric immunoglobulin containing an IgG1 hinge and IgG2b C_H2 and C_H3 domains. We predict that such an immunoglobulin would fix complement less effectively than a normal IgG2b molecule. Segmental flexibility is likely to be one factor correlated with complement fixation. The construction of immunoglobulin genes with novel combinations of domains or with site-specific mutations and their subsequent insertion into lymphoid cell lines should lead to deeper insights into how immunoglobulins mediate their distinctive functions. Methods for expressing designed immunoglobulins have recently been developed^{26,27}.

Monoclonal antibodies directed against tumour-specific antigens are now being used to treat certain malignancies²⁸. It will be important to learn whether the therapeutic effectiveness of a tumour-specific antibody depends on its isotype as well as on its antigen-binding properties. The capacity of such antibodies to interact optimally with tumour-specific antigens on cell surfaces may be influenced by their segmental flexibility.

We thank Tammy Guion for technical assistance and Drs Verne Schumaker and Juan Yguerabide for discussions. This research was supported by research grants from NIH (GM-24032 to L.S. and CA-04681, AI08917, and GM-17367 to L.A.H.). Fellowship support was provided by NIH (GM-07276 to T.M.V. and AI-06144 to V.T.O.), the Cystic Fibrosis Foundation (to J.R.) and the California Division of the American Cancer Society (J-22-82 to R.R.H.).

Received 22 August; accepted 5 October 1983.

- Noelken, M. E., Nelson, C. A., Buckley, C. E. & Tanford, C. *J. biol. Chem.* **240**, 218–224 (1965).
- Valentine, R. C. & Green, N. M. *J. molec. Biol.* **27**, 615–617 (1967).
- Feinstein, A. & Rowe, A. *J. Nature* **205**, 147–149 (1965).
- Marquart, M., Deisenhofer, J., Huber, R. & Palm, W. *J. molec. Biol.* **141**, 369–391 (1980).
- Amzel, L. M. & Poljak, R. J. *A. Rev. Biochem.* **48**, 961–967 (1979).
- Silverton, E. W., Navia, M. A. & Davies, D. M. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5140–5144 (1977).
- Yguerabide, J., Epstein, H. E. & Stryer, L. *J. molec. Biol.* **51**, 573–590 (1970).
- Cathou, R. E. *Compreh. Immun.* **5**, 37–83 (1978).
- Hanson, D. C., Yguerabide, J. & Schumaker, V. N. *Biochemistry* **20**, 6842–6852 (1981).
- Cebra, J. J. *et al. Proc. Robert A. Welch Foundation XVIII Immunochem.* Houston, Texas (ed. Milligan, W. O.) 67–96 (Welch Foundation, 1975).
- Dudich, E., Nezlín, R. S. & Franek, F. *FEBS Lett.* **89**, 89–92 (1978).
- Coffino, P., Laskov, R. & Scharff, M. D. *Science* **167**, 186–188 (1970).
- Liesegang, B., Radbruch, A. & Rajewsky, K. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3901–3905 (1978).

The scaled $\langle\phi\rangle$ value for IgE is $124 \times (150/190) = 98$ ns, which is even larger than the 81 ns value for IgG1. The sedimentation coefficient of 8.35 S for this IgE (V. Schumaker and M. Phillips, personal communication) corresponds to 6.59 S when scaled for the molecular weight difference between IgE and IgG. This is consistent with the time-averaged shape (as expressed by the translational diffusion coefficient) of IgE being very similar to that of the three IgG isotypes.

Reduction of the disulphide bonds between the heavy chains in the vicinity of the hinge between the Fab and Fc units^{16,24} provided additional information concerning the degree of segmental flexibility of the isotypes. The addition of 2 mM dithiothreitol (DTT) leads to a large decrease in $\langle\phi\rangle$ for IgE and IgG1 (Fig. 4B, Table 1). Reduction of the interchain disulphides lowers $\langle\phi\rangle$ of IgE from 124 to 88 ns, and that of IgG1 from 81 to 58 ns. DTT has no effect on $\langle\phi\rangle$ for the Fab fragment of IgG1. Thus, DTT lowers $\langle\phi\rangle$ of intact IgG1 by opening up a restricted hinge region. In contrast, DTT has almost no effect on $\langle\phi\rangle$ of IgG2a and IgG2b because the Fab units of these antibody molecules have a high degree of rotational freedom even before reduction of the interchain disulphides.

A large difference in segmental flexibility was also observed within the DNS3 family (Fig. 4C and Table 1). IgG2b is again much more flexible than IgG1, as evidenced by their $A(t)$ curves and $\langle\phi\rangle$ values. Also, DTT markedly decreases $\langle\phi\rangle$ of IgG1 and has little effect on $\langle\phi\rangle$ of the more flexible IgG2b (Fig. 4D and Table 1).

14. Dangel, J. L., Parks, D. R., Oi, V. T. & Herzenberg, L. A. *Cytometry* **2**, 395-401 (1982).
15. Stryer, L. *Science* **162**, 526-533 (1968).
16. Reidler, J. et al. *J. molec. Biol.* **158**, 739-746 (1982).
17. O'Farrell, P. Z., Goodman, H. M. & O'Farrell, P. H. *Cell* **12**, 1133-1142 (1977).
18. Oi, V. T., Bryan, V. M., Herzenberg, L. A. & Herzenberg, L. A. *J. exp. Med.* **151**, 1260-1274 (1980).
19. Hickman, S. & Kornfeld, S. *J. Immun.* **121**, 990-996 (1978).
20. Edelman, G. M. et al. *Proc. natn. Acad. Sci. U.S.A.* **63**, 78-85 (1969).
21. Kornfeld, R. & Kornfeld, S. A. *Rev. Biochem.* **45**, 217-237 (1976).
22. Yguerabide, J. *Meih. Enzym.* **26C**, 498-577 (1972).
23. Ishida, N., Ueda, S., Hayashida, H., Mitata, T. & Honjo, T. *EMBO J.* **1**, 1117-1123 (1982).
24. Chan, L. M. & Cathou, R. E. *J. molec. Biol.* **112**, 653-656 (1977).
25. Bengali, Z. H., Das, S. R. & Levine, P. H. *J. immun. Meih.* **33**, 63-00 (1980).
26. Oi, V. T., Morrison, S. L., Herzenberg, L. A. & Berg, P. *Proc. natn. Acad. Sci. U.S.A.* **80**, 825-829 (1983).
27. Rice, D. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7862-7865 (1982).
28. Miller, R. A. & Levy, R. *Lancet* **i**, 226-229 (1981).

LETTERS TO NATURE

Large anisotropy in the Universe does not prevent inflation

Marek Demianski*

Center for Relativity, Department of Physics,
The University of Texas at Austin, Austin, Texas 78712, USA

The inflationary model in which the Universe passes through an exponentially expanding phase can explain a few cosmological conundrums of the standard big-bang model. It is therefore interesting to study the stability of this model. It has been demonstrated that the inflationary scenario is stable with respect to density, velocity and gravitational wave perturbations¹⁻³. Here I show that the new inflationary model is compatible with large initial anisotropy. The initial anisotropic energy density could be many orders of magnitude larger than the initial radiation energy density.

The standard hot big-bang model of the Universe is remarkably simple and provides an amazingly accurate description of its evolution. There are, however, several observational facts which remain unexplained in this model, for example, the present high degree of isotropy of the microwave background radiation, the large scale homogeneity, and near critical energy density of the Universe.

The inflationary model of the early evolution of the Universe⁴⁻⁶ explains these facts without altering the basic predictions of the big bang model. Guth⁴ noticed that the first-order phase transition predicted by grand unified theories (GUTs), which occurs when temperature T is of the order of 10^{14} GeV, could have profound cosmological consequences. The inflationary scenario requires Higgs field effective potential $V(\phi)$ to have a global minimum at zero temperature. This global minimum is called the true vacuum. The zero temperature effective potential should also have a second metastable extremum at $\phi=0$ (the false vacuum). Furthermore, one assumes that there is a critical temperature $T_c \approx 10^{14}$ GeV above which the finite temperature effective potential has a lower value near the false vacuum than it has near the true vacuum. For any small temperature $T > 0$ the false vacuum is stabilized by a barrier in the finite temperature effective potential. The height of this barrier is of the order of T^4 and its width is of the order of T .

If at the very early stages of evolution the Universe contained a hot region ($T > 10^{14}$ GeV), which was isotropically expanding fast enough to cool to T_c before gravitational effects could cause it to collapse, then this region will cool to T_c and supercool. The constant false vacuum energy density determines then the rate of expansion and this region expands exponentially. This expansion can be described by the de Sitter metric with the false vacuum energy density having the role of a cosmological constant.

The Higgs field fluctuates around the false vacuum. Some fluctuations begin to grow and at some point become large enough to be described by a classical evolution equation. At the moment when evolution of fluctuations could be described by the classical equation the fluctuations should be small. If the

fluctuations grow very slowly, allowing many e folding times to pass, this single fluctuation region could grow to become larger than the present Universe. Finally, having reached the steep part of the potential, the field rapidly rolls down and oscillates about the true vacuum position. These oscillations are quickly damped out through coupling to other fields and the released energy is thermalized. The Universe reheats to temperature $T \approx T_c \approx 10^{14}$ GeV. The subsequent evolution follows the standard model.

The inflationary scenario can explain several cosmological conundrums, namely the isotropy of the microwave background radiation (the horizon problem), the observed large-scale homogeneity, the near critical energy density, and small number density of magnetic monopoles⁷.

It has been demonstrated that the inflationary scenario is stable with respect to inhomogeneous density, velocity, and gravitational wave perturbations¹⁻³. Hawking and Moss⁸ suggested that there is a cosmological 'no hair' theorem and the de Sitter space is dynamically stable. Barrow and Turner⁹ pointed out that in the inflationary model proposed by Guth⁴ large anisotropy prevents the de Sitter phase from occurring. I now show that the new inflationary scenario can tolerate very large homogeneous initial anisotropy.

Let us consider an anisotropic homogeneous Bianchi type 1 universe filled with radiation and constant false vacuum energy density ρ_v . The evolution of this model is described by

$$\dot{R}^2 = \frac{8\pi G}{3} (\rho_r + \rho_v) R^2 + \frac{8\pi G}{3} \frac{a^2}{R^4} \quad (1)$$

and

$$\rho_r R^4 = \text{constant} \quad (2)$$

where R is related to the volume expansion parameter by $\dot{V}/V = 3\dot{R}/R$, ρ_r is the radiation energy density, and a is a constant related to the initial anisotropy. Equation (1) can be rewritten in the form

$$\frac{dz}{d\tau} = 2 \left(z^2 + 1 + \frac{\beta}{z} \right)^{1/2} \quad (3)$$

where

$$\tau = \chi t = \left(\frac{8\pi G}{3} \rho_v \right)^{1/2} t$$

$$z = \left(\frac{R}{R_0} \right)^2 \left(\frac{\rho_v}{\rho_r(0)} \right)^{1/2}$$

and

$$\beta = \frac{a^2}{R_0^6 \rho_r(0)} \left(\frac{\rho_v}{\rho_r(0)} \right)^{1/2} = \frac{\rho_{an}(0)}{\rho_r(0)} \left(\frac{\rho_v}{\rho_r(0)} \right)^{1/2}$$

where R_0 is a constant, and $\rho_r(0)$ and $\rho_{an}(0)$ are the initial energy densities of radiation and anisotropy. In this equation z^2 is related to the false vacuum energy density, 1 to the radiation energy density, and β/z to the anisotropic energy density.

At the very early stage of expansion, when z is very small, the anisotropic energy density determines the expansion rate of the Universe, so that $z = \beta^{1/3} (3\tau)^{2/3}$. We are interested in a situation when the anisotropic energy density determines the expansion rate of the Universe at an epoch when the Universe

* Present address: Department of Physics and Astronomy, Williams College, Williamstown, Massachusetts 01267, USA. Permanent address: Institute of Theoretical Physics, University of Warsaw, Warsaw, Poland.

cools below T_c . When $T \approx T_c$ the radiation energy density and the vacuum energy density are equal, and $z = 1$. The anisotropic energy density will dominate the dynamics of expansion at that epoch if $\beta \gg 1$. Let us assume that the temperature of radiation at the Planck time $t_p = (G\hbar/c^5)^{1/2} = 0.5 \times 10^{-43}$ s is equal to Planck temperature $T_p = 1.2 \times 10^{19}$ GeV, then $\beta \gg 1$ if the initial anisotropic energy density $\rho_{an}(0) \gg 10^{10} \rho_r(0)$. The vacuum energy density starts to have a dynamically dominant role when $z = \beta^{1/3}$, ($\tau = \frac{1}{3}$). The de Sitter phase prevails if the small fluctuations of the Higgs field generated at the moment when or after the Universe cools to T_c ($\rho_v = \rho_r$, $z = 1$, $\tau = \frac{1}{3}(\beta)^{1/2}$) do not have enough time to grow significantly by the time when the vacuum energy density starts to determine the expansion rate of the Universe.

Let us assume that the initial fluctuations of the Higgs field are large enough to be described by the classical evolution equation

$$\ddot{\phi} + 3H\dot{\phi} = \lambda\phi^3 \quad (4)$$

where $\lambda \approx \frac{1}{2}$, and H is the Hubble constant, in our case $H = \frac{1}{3}t$. Assuming that initially, when $T = T_c$, $\phi(\tau = \frac{1}{3}(\beta)^{1/2}) = \phi_i$, we obtain an approximate solution of equation (4)

$$\phi = \frac{\phi_i}{\left(1 - \frac{\lambda}{2} \left(\frac{\phi_i}{\chi}\right)^2 \left(\tau^2 - \frac{1}{9\beta}\right)\right)^{1/2}} \quad (5)$$

Following Guth⁷ we assume that the initial fluctuations are of the order of thermal fluctuations and $\phi_i \approx 0.23\chi$. By the time that the vacuum energy density starts to determine the expansion rate of the Universe the fluctuations grow only insignificantly, $\phi(\tau = \frac{1}{3})/\phi_i - 1 \approx 7 \times 10^{-4}$. Therefore, the vacuum energy survives up to the moment when it starts to have a dominant dynamical role. The Universe enters the de Sitter phase and expands exponentially. The evolution of small fluctuations of the Higgs field in the de Sitter universe has been studied by Guth⁷. He showed that the fluctuations grow slowly and the Universe stays in the false vacuum state for many e folding times. The new inflationary model is, therefore, compatible with very large initial anisotropy. The initial anisotropic energy density can be many orders of magnitude larger than the initial radiation energy density.

One might ask what happens when due to coupling to other fields the anisotropic energy density is dissipated and released in the form of heat. This will only delay the moment of the phase transition but it does not change our main conclusion.

Recently, Futamase¹⁰ noticed that anisotropy will change the form of the effective potential $V(\phi)$. The additional terms are proportional to the anisotropic energy density and they distort the potential in a similar way as the high temperature corrections. Due to this additional bump in the potential the transition from the false to the true vacuum state would be slower, making the inflationary scenario even more plausible.

I thank the members of the Department of Physics of the University of Texas at Austin, Center for Relativity, and Center for Theoretical Physics for hospitality and support.

Correlation of bristlecone pine ring widths with atmospheric ^{14}C variations: a climate–Sun relation

C. P. Sonett

Department of Planetary Sciences and Lunar and Planetary Laboratory, University of Arizona Tucson, Arizona 85721, USA

H. E. Suess

Department of Chemistry, University of California at San Diego, La Jolla, California 93093, USA

Recent measurements by Stuiver and Quay¹ on wood growth during the past millennium confirm in an impressive way the contention that, in addition to being affected by the geomagnetic dipole moment, the production rate of ^{14}C is mediated by interplanetary modulation arising from solar activity, of which the sunspot index is a measure. But Stuiver², like others before him, finds no statistically significant correlation with climate indicators, including ring-width data and $\Delta^{14}\text{C}$. However, the data presented here show a case of an unusually convincing correlation between variations of cosmic ray-produced ^{14}C activity of the CO_2 in the terrestrial atmosphere during the past five millennia^{3,4} and annual growth ring widths for bristlecone pine wood from one particular area, that of Campito Mountain in eastern California.

The annual values of the Campito Mountain series were supplied by V. C. LaMarche of the University of Arizona Laboratory of Tree Ring Research. Twenty-year averages have been published by Lamb⁵. The bristlecone pine growth ring sequence was determined specifically to investigate the influence of climatic conditions on the growth rates of bristlecone pine trees. This growth ring sequence extends from 3405 BC to AD 1979. The sample trees are from the upper limit of the altitude band in which the bristlecone pine is found on Campito Mountain. For the high altitude (3,384 m) bristlecone pine at the upper growth region limit (from which the bristlecone data discussed here are drawn) the primary stress-induced growth effect is thought to be temperature^{6,7}.

The correlation of $\Delta^{14}\text{C}$ and tree-ring sequences is based on power spectral densities (PSDs) and cross-covariance; the PSDs are discussed first. The power spectrum of $\Delta^{14}\text{C}$ deduced from the La Jolla measurements of $\Delta^{14}\text{C}$ in samples of bristlecone pine wood from the White Mountains of California, has been published previously⁸. Essentially the same data are used here for a more detailed calculation of the $\Delta^{14}\text{C}$ spectrum. Additional comparisons (not shown here) have been made using the $\Delta^{14}\text{C}$ series reported by a workshop held at the University of Arizona⁹. A lesser correlation with a limber pine sequence from Mount San Geronio in Southern California has also been observed. Wood used for the ^{14}C determinations was preferentially selected from trees with abnormally wide rings.

The $\Delta^{14}\text{C}$ La Jolla measurements of bristlecone pine samples, dated dendrochronologically by Ferguson⁴, were used. They are unequally spaced in time, and therefore must be detrended and interpolated prior to computation. We first remove the quasi-sinusoid³ (thought to arise from secular drift in the intensity of the geomagnetic field), followed by application of a cubic spline¹⁰, quadratic detrending and finally interpolation at 10-yr intervals. To avoid extreme tensioning of the spline used for interpolation and yet cover all data points the highest frequency contributions are removed by a ± 20 yr moving linear least-squares estimator which replaces each interpolated point by a value which dampens the highest frequency components. The LaMarche bristlecone pine ring-width sequence, consisting of yearly data points, is filtered and decimated by 10; the initial computations were made using the sequences truncated to the

Received 8 July; accepted 7 October 1983.

1. Frieman, J. A. & Will, C. M. *Astrophys. J.* **259**, 437–441 (1982).
2. Barrow, J. D. in *The Very Early Universe* (eds Gibbons, G., Hawking, S. & Siklos, S.) (Cambridge University Press, 1983).
3. Boucher, P. & Gibbons, G. in *The Very Early Universe* (eds Gibbons, G., Hawking, S. & Siklos, S.) (Cambridge University Press, 1983).
4. Guth, A. H. *Phys. Rev. D* **23**, 347–356 (1981).
5. Linde, A. D. *Phys. Lett.* **108B**, 389–395 (1982).
6. Albrecht, A. & Steinhardt, P. J. *Phys. Rev. Lett.* **48**, 1220–1223 (1982).
7. Guth, A. H. in *The Very Early Universe* (eds Gibbons, G., Hawking, S. & Siklos, S.) (Cambridge University Press, 1983).
8. Hawking, S. & Moss, I. G. *Phys. Lett.* **110B**, 35–38 (1982).
9. Barrow, J. D. & Turner, M. S. *Nature* **292**, 35–38 (1982).
10. Futamase, T. *Coleman-Weinberg Breaking in an Anisotropic Space-time* (Preprint, University College Cardiff, 1983).

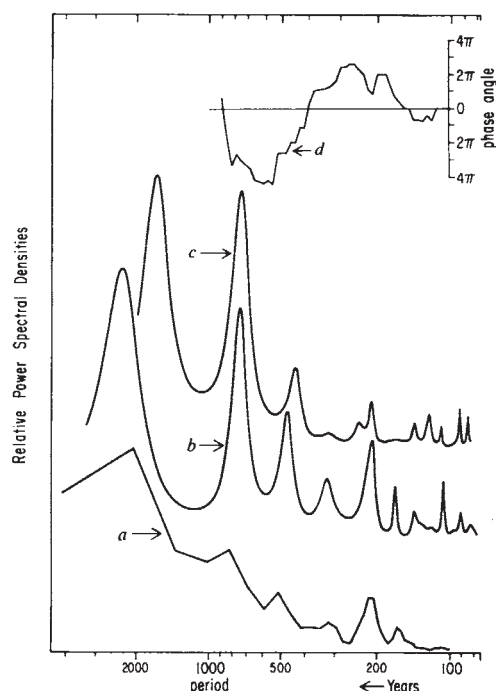


Fig. 1 Power spectral densities of the La Jolla $\Delta^{14}\text{C}$ and bristlecone pine tree-ring growth record: *a*, the Neftel *et al.*⁸ FFT of the full 8,400-yr record; *b*, the maximum entropy (MEM) spectrum at autoregressive (AR) order 120 of the truncated 5,290-yr (3405 BC-AD 1885) La Jolla record; *c*, the Campito Mountain bristlecone pine (3405 BC-AD 1885) growth ring spectrum at AR order 120; *d*, variation in relative phase between $\Delta^{14}\text{C}$ and Campito Mountain bristlecone pine ring growth with frequency and period. The phase is that corresponding to the peak coherence between narrow-banded versions of the original sequences obtained from Fourier spectra filtered to isolate narrow spectral intervals which provide the basis for narrow band $\Delta^{14}\text{C}$ and bristlecone pine sequences used for correlation.

maximum common time period (3405 BC-AD 1885; $T = 5,290$ yr).

As the primary ^{14}C spectrum is 'red' with a slope, using a power law fit, with spectral index approximately -1 , and the region of the spectrum of interest here encompasses periods longer than about 80 yr, aliasing is not significant. Filtering and corruption by splining have been studied using Monte Carlo tests and synthetic sequences whose spectra satisfy the 'red' slope characterizing the $\Delta^{14}\text{C}$ spectrum. The test sequence realizations are generated using sinusoidal sequences of decreasing mean amplitude, normally distributed, and with random phase¹¹, upon which may be superimposed signals of specific intensity. These tests show insignificant corruption of the data for periods in excess of several times 10 yr.

The PSDs are computed using the maximum entropy method (MEM)^{12,13}, although we have confirmed major spectral features using the less constrained Yule-Walker (YW) algorithm^{14,15} and also the Singleton fast Fourier transform (FFT)¹⁶. Both the MEM and YW algorithms involve all-pole filter models which are considered most suitable for data where the major features are spectral lines. Figure 1 shows the PSDs for both the La Jolla $\Delta^{14}\text{C}$ and the Campito Mountain bristlecone pine tree-ring sequences. Both have major features in common. As noted the periods of the lines should be regarded as approximate, because the pole positions in the autoregressive (AR) filter are somewhat dependent upon detailed properties of the sequences. (This is thought to be due to the line splitting phenomenon which results in peak shifting depending upon signal to noise ratio¹⁷. Its correction is under study and is expected to refine the periods reported here.) The La Jolla $\Delta^{14}\text{C}$ series has been previously investigated by Kruse³ and in greater detail by Neftel *et al.*⁸ using the standard FFT. The prime purpose of these investigations was to demonstrate that the

secular $\Delta^{14}\text{C}$ variations are not random and hence cannot be caused by experimental errors. Moreover these investigations show that the $\Delta^{14}\text{C}$ time series contains cyclic features, in particular a prominent spectral line of about 200-yr period. Our analysis confirms these earlier results in all their details. Because of its ubiquity, we have centred attention on the 200-yr period as a key marker of a solar interrelation between the sequences reported here. Thus the appearance of the 200-yr period in the ^{14}C record, repeated in the tree-ring record, strongly supports the contention that this solar-based signal is reflected in the growth rate of wood in the bristlecone pines.

The PSDs have normalized amplitudes with the highest peak in each spectrum set to unity for convenience. Table 1 gives the periods of selected lines for each of the two spectra. Although MEM estimates are thought to have better intrinsic resolution than $\Delta f = 1/T$, FFT estimates are conservatively rated at $\Delta f > 6/T$. The actual MEM resolution is possibly less than that given above, since MEM carries intrinsic filtering depending inversely upon order; very approximately the MEM frequency resolution is optimistic by perhaps a factor of 3-4. This asymptotic limit is also given by Ulrych and Bishop¹³ based on an analysis of Kromer¹⁸. Thus $\Delta f_r = 2.3 \times 10^{-4} \text{ yr}^{-1}$ for the bristlecone sequence and $\Delta f_c = 1.9 \times 10^{-4} \text{ yr}^{-1}$ for La Jolla, assuming no filtering so that (0) $T/20$ independent degrees of freedom are available. Combining the two gives an approximate value of $\Delta f_s = 4.2 \times 10^{-4} \text{ yr}^{-1}$, but more probably greater than 10^{-4} yr^{-1} . Inspection of Table 1 discloses that of the lines tabulated, three pairs differ in frequency by 10^{-4} yr^{-1} , one pair by $3 \times 10^{-4} \text{ yr}^{-1}$ and the two high frequency pairs by $7-8 \times 10^{-4} \text{ yr}^{-1}$. Thus a majority of the line pairs tabulated lie approximately within or better than the estimated resolution, and the two exceptions are within two resolution elements.

The calculations reported here have been repeated for the $\Delta^{14}\text{C}$ sequence from the Arizona workshop, but for brevity we report only that the agreement is qualitatively reasonable. As a final test we have also compared the La Jolla $\Delta^{14}\text{C}$ sequence with a tree-ring growth sequence of limber pine from Mount San Geronio in Southern California. The 100- and 200-yr lines match closely; those at longer periods (including the very strong line at 500 yr in the limber pine spectrum) probably correspond to the two long periods in the two ^{14}C spectra.

Note also that these time sequences are statistically non-stationary. For a certain period within the full La Jolla record the 200-yr line disappears as has been pointed out previously by Neftel *et al.*⁸, although the 200-yr line is strong before and after this period. Thus the record used for the work reported here should be regarded as an integrated composite of spectral response over the entire time period examined.

An independent verification of the existence of a correlation between the La Jolla $\Delta^{14}\text{C}$ sequence and the BCP ring growth

Table 1 ^{14}C data for La Jolla (3405 BC-AD 1885) (*a*) and Campito Mountain bristlecone (3405 BC-AD 1885) (*b*)

	Frequency (yr^{-1})	Period T (yr)	$T - \Delta T$	$T + \Delta T$
<i>a</i>	0.0004	2,500	1,700	4,740
	0.0014	714	629	826
	0.0047	213	205	222
	0.0087	115	113	118
	0.0096	104	102	106
	0.0114	88	86	89
<i>b</i>	0.0007	1,430	1,125	1,957
	0.0015	667	592	763
	0.0048	208	200	217
	0.0088	114	111	116
	0.0103	97	95	99
	0.0122	82	81	83

Frequency and period of some selected major lines in the ^{14}C and bristlecone ring growth spectra. The two right-hand columns are the period resolution, estimated from the sequence length.

sequence is obtained by a narrow band moving window cross-correlator operating jointly on the two sequences. The two time domain sequences are first Fourier transformed, yielding the complex Fourier coefficients. The fundamental resolution in frequency is defined as follows: as the length of the sequences, $T = 5,290$ yr, the frequency resolution, $\Delta f = 1/5,290 \text{ yr}^{-1}$. The Nyquist frequency, f_N , is half the final adjusted sampling frequency of 10 yr^{-1} , so the number of spectral estimates $n = f_N/T$. We apply a simple narrow-band filter consisting of a linearly tapered window of full width $7/20 \text{ yr}^{-1}$. This triangular window provides imperfect but reasonable suppression of side lobes. Since there are five non-zero weights ($1/3, 2/3, 1, 2/3, 1/3$) in this filter, it weights the ends of the sequence asymmetrically. For this reason application of the filter is restricted to those spectral estimates lying in the frequency interval $1/1,770 < f < 1/90 \text{ yr}^{-1}$. This eliminates the two lowest and highest frequency estimates.

This procedure yields a set of spectra encompassing only a very narrow band of frequencies in accordance with the discussion just given. Inversion of these spectra into the time domain thus provides two subsets of sequences, one for $\Delta^{14}\text{C}$ and the other for bristlecone pine growth rings. We now correlate these in pairs corresponding to each frequency estimates. Such a narrow banded correlation exhibits time lags which may be interpreted as phase lags over narrow frequency intervals. Further, these lags are relatively independent of the coherence. We now choose those lags corresponding to the peak value of coherence for each of the narrow-band sequence pairs. The correspondence between lag and phase with varying frequency yields the phase-frequency relation shown in fig. 1. This figure indicates a continuously varying phase with increasing frequency. The variation in phase exceeds that which could be assigned to reservoir effects, those being limited to a lag of $\pi/2$.

Stuiver², for the AD centuries investigated by him, reports spectral features in his ^{14}C production rate curves for the periods $T = 126\text{--}141$ and $185\text{--}218 \text{ yr}$ at the 2σ confidence level or better. Assuming his smoothed FFTs have $\Delta f = 1/15 \text{ yr}^{-1}$, then the 200-yr line is detected; the 120-yr line is less certain, lying slightly outside his lower period window. However, the correlation (of very low confidence) between White Mountain trees and the $\Delta^{14}\text{C}$ record exhibits zero lag. Since coherence depends on bandwidth, the independent test for existence of a meaningful relative phase which we use is preferred. The credibility of a detectable phase is checked independently of MEM though supported by calculations using a two channel cross-spectral MEM estimator. We believe that further work is required to interpret the phase relation, but it can be said that its detection is qualitatively consistent with the comparison of the $\Delta^{14}\text{C}$ and BPC ring growth auto-spectra.

The relation between $\Delta^{14}\text{C}$ and solar activity is well established, though many details are poorly understood. The detection of spectral lines in the bristlecone pine growth ring spectra at periods commensurate with those in the $\Delta^{14}\text{C}$ spectra implies that the $\Delta^{14}\text{C}$ detection of solar signals is repeated in the ring growth of the Campito Mountain bristlecone pines.

The similarity between the $\Delta^{14}\text{C}$ record and the growth record of the Campito Mountain bristlecone pine suggests that a forcing function is present which mediates both the atmospheric ^{14}C level and tree growth. The expectation is that the ^{14}C uptake by trees will lag growth on the premise that growth is a short term response to climate, whereas atmospheric and oceanic reservoirs contribute a frequency dependent lag to the net atmospheric ^{14}C budget¹⁹. Since the two effects discussed here appear correlated, by a process of elimination, the Sun is an obvious candidate for the source of the variations, though expressed through widely differing 'channels', that is, hydromagnetic in the case of ^{14}C and an as yet poorly understood mechanism in the case of tree growth. Further work, especially in regard to the very complex problem of the relative phase, may aid in clarifying the mechanism(s).

As there is a strong stochastic element in the data used for this analysis, the results are fundamentally probabilistic in nature

and cannot be assigned a numerically deterministic certainty. This also has the effect that the actual signal strength, for what are generally to be regarded as threshold signals, has a probabilistic distribution and can attain larger or smaller values than measured, with the mean favouring the larger. Corrections for this effect have not been considered here but would have the general effect of increasing the true signal levels above those inferred from measurement²⁰. Lastly, because phase is much more sensitive to noise corruption than is amplitude, details of the relative phase should be regarded as tentative.

We thank C. W. Stockton for providing the bristlecone and limber pine tree ring width sequences, T. W. Linick for the La Jolla $\Delta^{14}\text{C}$ data and to P. E. Damon for providing the workshop sequence. Programming was carried out by T. Trebisky and by L. D. Smith. CPS was supported by grant ATM-8308378 from the NSF Solar-Terrestrial Program.

Received 25 April; accepted 18 October 1983.

1. Stuiver, M. & Quay, P. D. *Science* **207**, 11 (1980).
2. Stuiver, M. *Nature* **286**, 868 (1980).
3. Suess, H. E. *Radiocarbon* **20**, 1-8 (1978), **22**, 200-209 (1980).
4. Linick, T. W., in preparation.
5. Lamb, H. H., *Climate, Present, Past, and Future* Vol. 2, 59 (Methuen, London, 1977).
6. Fritts, H. C. *Tree Rings and Climate*, (Academic, New York, 1976).
7. LaMarche, V. C. *Science* **183**, 1043 (1974).
8. Neftel, A., Oeschger, M. & Suess, H. E. *Earth planet. Sci. Lett.*, **56**, 127 (1981).
9. Klein, J., Lerman, J. C., Damon, P. E. & Ralph, E. K. *Radiocarbon*, **24**, 103 (1982).
10. Reinsch, C. H. *Num. Math.*, **10**, 177 (1967).
11. Rice, S. O., reprinted from Bell System Tech. Jour. in *Selected Papers on Noise and Stochastic Processes* (ed. Wax, N. Dover, 1954).
12. Burg, J. P. thesis, Stanford Univ. (1975).
13. Ulrych, T. J. and Bishop, T. N. *Rev. Geophys. Space Phys.* **13**, 183 (1975).
14. Yule, U. Y., *Proc. R. Soc. A* **642**, 267 (1927).
15. Box, G. E. P. and Jenkins, G. M. *Times Series Analysis* (Holden-Day 1976).
16. Singleton, R. C. *IEEE Trans. Audio Electroacoust.* **17**, 93 (1969).
17. Jaynes, E. T. *Proc. IEEE* **70**, 939 (1982).
18. Kromer, R. thesis, Stanford Univ. (1970).
19. Houtermans, J. C., Oeschger, M. & Suess, H. E. *J. geophys. Res.* **78**, 1893-1908 (1973).
20. Groth, E. J. *Astrophys. J. Suppl. Ser.* **286**, 29, 285 (1975).

Reconstruction of precipitation history in North American corn belt using tree rings

T. J. Blasing & Daniel Duvick

Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830, USA

Tree-ring-based reconstructions of past climate have proved useful for evaluating hydrological statistics calculated from the shorter observational record of climate¹, and particularly for analysing past patterns of drought in the western USA². The independently verified climatic reconstructions obtained here provide a significant contribution to a reliable data base for evaluation of climatic changes on time scales of decades to centuries. Fifteen tree-ring chronologies were developed from specimens of drought-sensitive white oak (*Quercus alba*), many over 300-yr old, in the western part of the North American corn belt. Verification statistics indicate reliable reconstructions of annual precipitation for the two-state area of Iowa and Illinois can be obtained back to 1680 from these chronologies. This extends the network of confirmed drought-sensitive tree-ring chronologies well into the central USA, and provides encouragement for further expansion of the tree-ring chronology network into the North American grain-belt region.

Tree rings have been used to reconstruct past climate in several regions including the western USA^{2,3}, western Canada and Alaska^{4,5}, the eastern USA^{6,7}, and the British Isles⁸. Stockton and Meko² reconstructed a 270-yr drought history of the western USA, which contained some evidence of a 22-yr Hale sunspot cycle⁹. Their study region extended into the central USA, but their tree-ring data base was weak there, including only a few sites on the periphery of the North American grain-belt region. Furthermore, the climatic information content of

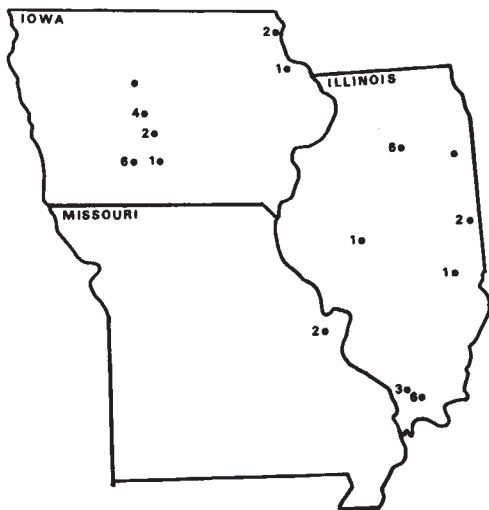


Fig. 1 Site locations of trees and number of trees at each location which date back to at least AD 1680. One site in Iowa and one site in Illinois had no trees that old, but all sites had at least one tree dating back to 1686. Some younger trees from each of the sites were also used in the reconstruction.

the ring-width chronologies from the central United States had not been verified.

Since that study appeared², we have published¹⁰ preliminary results, in which the state average annual precipitation for Iowa was reconstructed back to 1680 using three tree-ring-width chronologies from the central part of the state. Our verification results encouraged further investigations leading to the more extensive network of tree sites used in this study (Fig. 1).

Usually two cores, each 4 mm in diameter, were extracted at breast height from opposite sides of each tree sampled. Ring widths from each core were converted to ring-width indices¹¹ using spline functions to prefilter non-climatic variance due to gradual changes in tree age, competitive status, and so on¹². Ring-width indices of all cores at all sites in each state, Iowa and Illinois, were averaged for each year (with the Missouri site data included in the Illinois state average), and the two statewide series were averaged to form a regional ring-width-index series. This was used as the predictor variable in linear regression to estimate annual precipitation amounts averaged for the two-state area (Iowa and Illinois). In this study, annual precipitation was computed from August of the preceding year through July of the year corresponding to ring formation. This coincides with the 12-month period for which these tree rings best reflect total precipitation amounts, from the end of the preceding growing season to the end of the season in which the ring is formed¹⁰. Ring-width data from 1920 to 1980, and the corresponding annual precipitation data as defined above, were used to calculate the regression coefficients. State average-precipitation data available for both states since 1878¹³ were used for independent verification of precipitation estimates up to 1919.

The correlation coefficient between actual and estimated precipitation values was +0.79 for the period 1920–80, and +0.86

for the independent period. For comparison, when annual (August–July) precipitation data at Springfield, Illinois, and Des Moines, Iowa, were averaged, the correlation between that series and the two-state averages for the period 1920–80 was +0.89. All of these are significant well beyond the 99% confidence level, even after first-order autocorrelations (always below +0.25) were accounted for¹⁴. For the independent data (1878–1919) the reduction of error, RE, was +0.66. The RE is described by Fritts *et al.*³ as “a most rigorous verification statistic”. It is, in some ways, analogous to the so-called explained variance, but its range extends from $-\infty$ to +1, and any positive value indicates some information in the reconstruction³. Precipitation is often noticeably underestimated for extremely wet years, presumably because moisture is not limiting to growth¹⁰. Thus, precipitation estimates for normal and dry years are often more accurate than the overall statistics indicate. In this case, the RE increased to +0.76, although the correlation coefficient remained at +0.86, when one-third of the years (14 yr) having the highest estimated precipitation values were removed from the independent data set. The overall accuracy and reliability of these reconstructions of past precipitation amounts, as measured by the above statistics, are as great if not greater than any reported to date.

The reconstructed annual precipitation since 1680 is shown in Fig. 2a. No droughts appreciably worse than the ‘dust bowl’ of the 1930s are indicated, although there apparently were several droughts of roughly equal severity. The reconstructions show no long-term changes such as might occur if the precipitation there were linked to Northern Hemisphere temperature trends¹⁵. This is also true of the (unfiltered) actual precipitation series which began in 1878 (Fig. 2a). Droughty conditions are apparent beginning in the mid 1880s when the Northern Hemisphere was relatively cool¹⁵, and in 1816 when the ‘year without a summer’ occurred in the northeastern United States and adjoining Canadian provinces¹⁶. However, the ‘dust bowl’ drought of the 1930s coincided with one of the warmest periods in the last century and probably one of the warmest periods in the span of our reconstructions which extend about two centuries back into the Little Ice Age^{17,18}. Overall, drought in the western corn belt appears to have been about equally probable during the relatively warm and cool periods of the past three centuries. With acknowledgment to projections of increased droughtiness for a CO₂-induced climatic warming¹⁹, our reconstructions indicate that droughts comparable with those of the 1890s and 1930s afflicted the Iowa–Illinois region at various times over the past three centuries and, therefore, should be expected to continue to do so even if no further warming occurs.

Stockton and Meko²⁰ analysed tree-ring data from our sites in central Iowa and found evidence of a 22-yr cycle. However, the cycle was not persistent throughout the period of record. It seemed to be strongest since about 1890 and was not evident before 1800. The 10-yr moving averages in Fig. 2b, based on

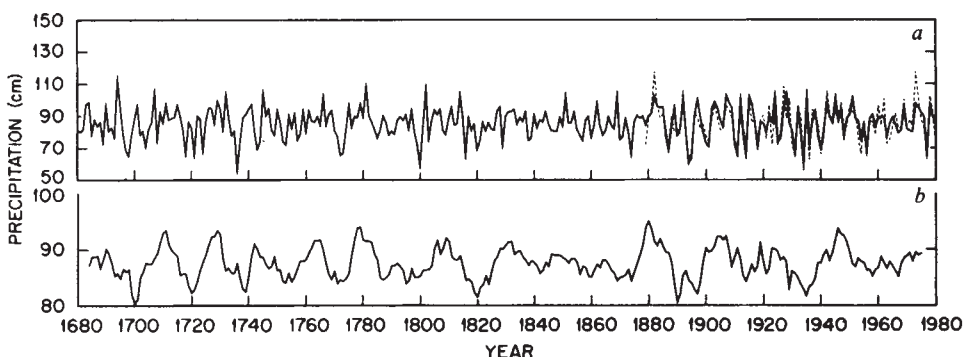


Fig. 2 a, Reconstructed annual (August–July) precipitation for Iowa–Illinois as obtained from the regression equation: $y = -2.788 + 90.756x$, where y is the precipitation estimate for a particular year and x is the regional ring-width index for the corresponding year. Actual values including independent data are indicated by the dashed line. When this equation, derived using data from young and old trees, was applied to reconstruct precipitation from only the cores dating back to 1680, the RE statistic corresponding to +0.66 in the text changed negligibly to +0.68. b, The 10-yr running means of the reconstructions plotted for the fifth year of each 10-yr sequence.

the data now available, suggest a cycle of about 20-yr before 1800. That may be at least partially an artefact of the frequency-response function of the moving average, but the reconstructed drought around 1720, for example, in Fig. 2b is much more pronounced than in the 10-yr moving averages presented earlier¹⁰ which were based only on chronologies from central Iowa. It has been suggested that a tendency for drought to occur about every 20 yr may be due to combined effects of the Hale solar cycle of about 22 yr and the regular lunar-tidal cycle of 18.6 yr (refs 21, 22).

Reconstructions of past climate provide an extended data base for formulating or testing hypotheses about possible causes of climatic change (such as solar and lunar cycles), and for evaluating or recalculating probabilities of climatic events (such as drought) calculated from the shorter instrumental records. This work shows that excellent reconstructions of annual precipitation, for at least three centuries, are possible for portions of the North American grain belts.

We thank Dale Huff and John Krummel for helpful comments. The research was sponsored by the NSF, Ecosystem Studies Program, under Interagency Agreement DEB81-15316 with the US Department of Energy under contract W-7405-eng-26 with Union Carbide Corporation. Publication no. 2243, Environmental Sciences Division, ORNL.

Received 6 September; accepted 4 October 1983.

1. Stockton, C. W. *Long-Term Streamflow Records Reconstructed from Tree Rings* (University of Arizona Press, Tucson, 1975).
2. Stockton, C. W. & Meko, D. M. *Weatherwise* **28**, 244-249 (1975).
3. Fritts, H. C., Lofgren, G. R. & Gordon, G. A. *Quat. Res.* **12**, 18-46 (1979).
4. Garfinkel, H. L. & Brubaker, L. B. *Nature* **286**, 872-873 (1980).
5. Jacoby, G. C. & Cook, E. R. *Arctic Alp. Res.* **13**, 409-418 (1981).
6. Cook, E. R. & Jacoby, G. C. *Science* **198**, 399-401 (1977).
7. Cleaveland, M. K. thesis, Clemson Univ., South Carolina (1975).
8. Briffa, K. R., Jones, P. D., Wigley, T. M. L., Pilcher, J. R. & Baillie, M. G. L. *J. Climatol.* (in the press).
9. Mitchell, J. M., Stockton, C. W. & Meko, D. M. in *Solar-Terrestrial Influences on Weather and Climate* (eds McCormac, B. M. & Seliga, T. A.) 125-143 (Reidel, Dordrecht, 1979).
10. Duval, D. N. & Blasing, T. J. *Wat. Resour. Res.* **17**, 1183-1189 (1981).
11. Fritts, H. C. *Tree Rings and Climate* (Academic, London 1976).
12. Cook, E. R. & Peters, K. *Tree-Ring Bull.* **41**, 45-53 (1981).
13. U.S. Weather Bureau (now, National Weather Service). *Climatological Data for the United States by Sections*. (National Climatic Center, Asheville, various dates).
14. Bartlett, M. S. *J. R. statist. Soc.* **98**, 536-543 (1935).
15. Jones, P. D., Wigley, T. M. L. & Kelly, P. M. *Mon. Weath. Rev.* **110**, 59-70 (1982).
16. Hughes, P. *Weatherwise* **32**, 108-111 (1979).
17. Groveman, B. S. & Landsberg, H. Publ. 79-181 (Meteorology Program, Univ. of Maryland, 1979).
18. Bryson, R. A. & Murray, T. J. *Climates of Hunger* (University of Wisconsin Press, Madison, 1977).
19. Manabe, S., Wetherald, R. T. & Stouffer, R. J. *Climat. Change* **3**, 347-386 (1981).
20. Stockton, C. W. & Meko, D. M. *J. Climatol. appl. Met.* **22**, 17-29 (1983).
21. Vines, R. G. *J. geophys. Res.* **87**, 7303-7311 (1982).
22. Gough, D. *Nature* **288**, 639-640 (1980).

Elemental constituents of atmospheric particulates and particle density

Akiyoshi Sugimae

Environmental Pollution Control Center, Osaka Prefecture, 3-62, 1-chome, Nakamichi, Higashinari-ku, Osaka 537, Japan

Atmospheric particulates are expected to be formed from materials of various densities. Particle motion in the atmosphere is characterized by the product of the density multiplied by the diameter squared. For studies of source characteristics, atmospheric transport processes and removal rates of atmospheric particulates, accurate knowledge of density distribution is essential. Previous studies¹⁻³ indicated that the value of 1.7 g cm^{-3} for the average density of atmospheric submicrometre particles was quite reasonable. However, there are very few published data on the elemental constituent as a function of density of atmospheric particulates. I report here the density distribution of Si, Fe, Mn, Zn and Cu in the atmosphere.

Atmospheric particulates were sampled either by filtration from the suspension in the atmosphere, or by collection of

deposited particles as they fall out under gravitational influence, known as dustfall. Samples of suspended particulates were collected on the rooftop (about 10 m above street level) of the Environmental Pollution Control Center in a metropolitan area of Osaka, using both a conventional high-volume sampler (hi-vol) and a low-volume sampler (low-vol). The hi-vol was used to collect samples of total suspended particulates (TSPs) on 8×10 -inch filters. The sampler is fitted in an aluminium shelter which prevents direct settling of dusts onto the filter. Particles with a terminal settling velocity in air greater than approximately 30 cm s^{-1} are excluded from the shelter⁴. The maximum theoretical particle size has been calculated to be $100 \mu\text{m}$ (particle density = 1.0 g cm^{-3}). However, collection efficiencies for large particles are quite sensitive to small changes in wind velocity, approaching angle and flow rate⁵. The upper limit of particle sizes actually collected is believed to be $15\text{--}20 \mu\text{m}$. The Japanese Environment Agency specifies the use of low-vol for the long-term monitoring of inhalable particulate (IP) levels. Coarse particles are preliminarily separated by three cyclones installed in the sampler. The low-vol has a D_{50} (aerodynamic diameter at an effectiveness of 50%) = $8 \mu\text{m}$ and collects IPs on 48-mm diameter filter at a flow rate of 20 l min^{-1} .

The polystyrene filter (Microsorban type 99/97, Delbag-Luftfilter GMBH, Berlin) was used as the collection medium, because the filter has collection efficiencies, capacities and loading characteristics similar to the more commonly used glass fibre filter, yet has a much lower blank concentrations of trace elements⁶.

Particulates which fell as a result of their own settling velocity and which were carried down with precipitation were collected in a deposit gauge. The sample collection continued for a month. After collection, water-insoluble component of dustfall was filtered off with distilled water on a polystyrene filter.

To evaluate the contribution of wind-blown soil to local atmospheric particulates, a soil sample was taken near the sampling station.

I have modified the density gradient separation technique used for the chemical speciation of Pb compounds in street dusts⁷⁻¹⁰ and for the multi-element characterization of roadway dusts¹¹ and applied it to density fractionation of particulate samples collected on polystyrene filters. The use of this type of filter was a good choice because a suspension of particulate samples for the subsequent density separation was prepared by dissolving the filter material in dichloromethane. After centrifugation, the float was transferred and the residue was centrifuged in a series of dichloromethane-diiodomethane mixtures such that the density of each successive mixture was incrementally increased by 0.2 g cm^{-3} . The final centrifugation was in diiodomethane. Particulate samples were thus classified into 12 density fractions from <1.3 to $>3.3 \text{ g cm}^{-3}$. Iron, Mn, Zn and Cu were determined in each density fraction by inductively coupled plasma-atomic emission spectrometry (ICP-AES) following acid digestion in hot aqua regia. Silicon was determined by ICP-AES after alkaline fusion. The experimental methods have been described in detail elsewhere¹²⁻¹⁵.

The densities of the commonly recognized compounds which exist or are expected to exist in the atmosphere are as follows: SiO_2 (2.65 g cm^{-3}), Fe_2O_3 (5.24 g cm^{-3}), Fe_3O_4 (5.18 g cm^{-3}), $\text{Fe}_2(\text{SO}_4)_3$ (3.094 g cm^{-3}), MnO (5.026 g cm^{-3}), ZnO (5.61 g cm^{-3}), ZnSO_4 (3.54 g cm^{-3}), CuO (6.0 g cm^{-3}) and CuSO_4 (3.61 g cm^{-3}).

Commonsense suggests that heavy metal compounds will be present primarily associated with particles of high density ($>3.3 \text{ g cm}^{-3}$). However, the particle density is often related to the mode of origin, composition and interaction with the surrounding vapours and with each others¹⁶. The density of particle formed by comminution, attrition or disintegration will resemble that of parent material. On the other hand, fine particles originating from nucleation of vapours are highly agglomerated. It is generally accepted that the density of agglomerate which includes void spaces is less than that of individual component particles.

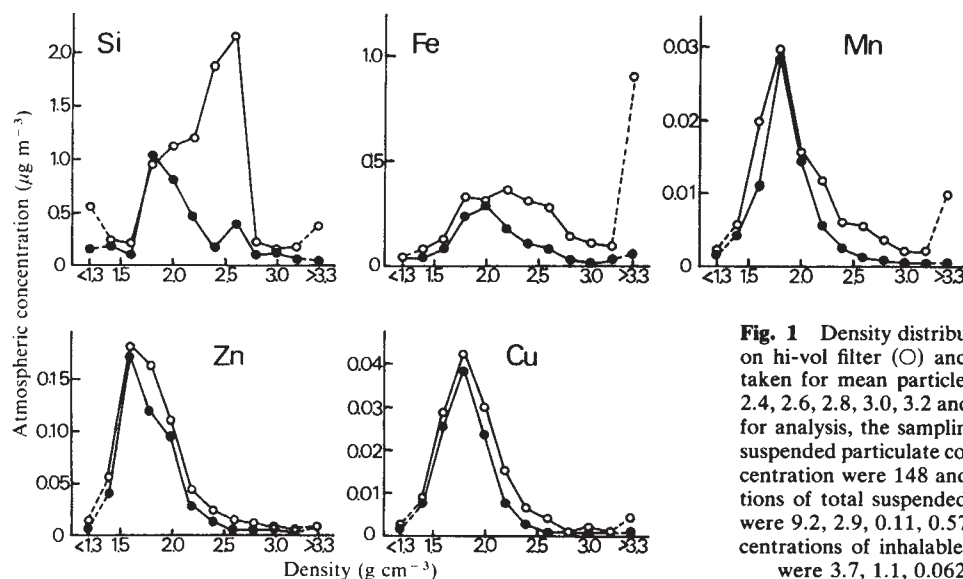


Fig. 1 Density distributions for Si, Fe, Mn, Zn and Cu collected on hi-vol filter (○) and low-vol filter (●). Measurements were taken for mean particle density of <1.3, 1.4, 1.6, 1.8, 2.0, 2.2, 2.4, 2.6, 2.8, 3.0, 3.2 and >3.3 g cm⁻³. To collect enough material for analysis, the sampling duration was extended to 100 h. Total suspended particulate concentration and inhalable particulate concentration were 148 and 69 μg m⁻³, respectively. The concentrations of total suspended particulates for Si, Fe, Mn, Zn and Cu were 9.2, 2.9, 0.11, 0.57 and 0.14 μg m⁻³, respectively. The concentrations of inhalable particulates for Si, Fe, Mn, Zn and Cu were 3.7, 1.1, 0.062, 0.46 and 0.11 μg m⁻³, respectively.

The elemental density distributions shown in Fig. 1 represent that the concentration of Si in TSPs is relatively high in the density range of 2.5–2.7 g cm⁻³, corresponding to the density of SiO₂. Approximately 30% of Fe particles in TSPs collected by means of hi-vol are associated with particles of density higher than 3.3 g cm⁻³. There is a general tendency for the soil-derived elements in TSPs, typified by Si and Fe, to increase in concentrations in the corresponding density ranges of the parent materials (such as quartz; haematite and magnetite). On the other hand, the concentration of Si in IPs goes through a maximum in the density range of 1.7–1.9 g cm⁻³ and another small maximum in the density range of 2.5–2.7 g cm⁻³. The density distribution of inhalable Fe particles shows definite peak in the density range of 1.9–2.1 g cm⁻³. Iron particles with density higher than 3.3 g cm⁻³ can account for only a few per cent of the total mass of Fe particles. It is noticeable that TSPs and IPs display quite different density distributions with respect to Si and Fe particles, which are generally present as coarse particles in suspended particulates^{17–24}. The low-vol used for this study is limited to the collection of particles which have aerodynamic diameters <8 μm. The results suggest that the low-vol is relatively less efficient for collecting atmospheric soil particles existing as coarse particles in the atmosphere.

The presence of Si and Fe particles with densities less than those of parent materials is attributable partially to the presence

of soil agglomerates with low apparent densities^{25,26} and partially to the presence of empty spheres, known as cenospheres and plerospheres^{27–29}; both spheres are formed from a particular combustion.

Volatile (chalcophilic) elements such as Zn and Cu are concentrated in low-density fractions which seem to consist of carbonaceous matter judging from their black colour and microscopic observation. In the density distributions of Zn and Cu, only one peak is observed in the density range of 1.5–1.7 and 1.7–1.9 g cm⁻³, respectively. It is also notable that the density distributions of Zn and Cu do not differ appreciably between hi-vol and low-vol sample. The densities of Zn and Cu compounds are anticipated to be higher than 3.3 g cm⁻³ because the likely compounds of these elements in the atmosphere have the bulk densities, being higher than 3.3 g cm⁻³. Nevertheless, Zn and Cu particles with the density higher than 3.3 g cm⁻³ can account for only a few per cent of total mass of Zn and Cu, respectively. This is probably because these elements are capable of being volatilized in high-temperature, anthropogenic processes. And fine particles originating from volatilization grow by condensation and coagulation to form agglomerates with low densities.

In the case of Mn particles collected on hi-vol filter, the density distribution shows two maxima, one for particles with density of the 1.7–1.9 g cm⁻³ range and the other for particles with

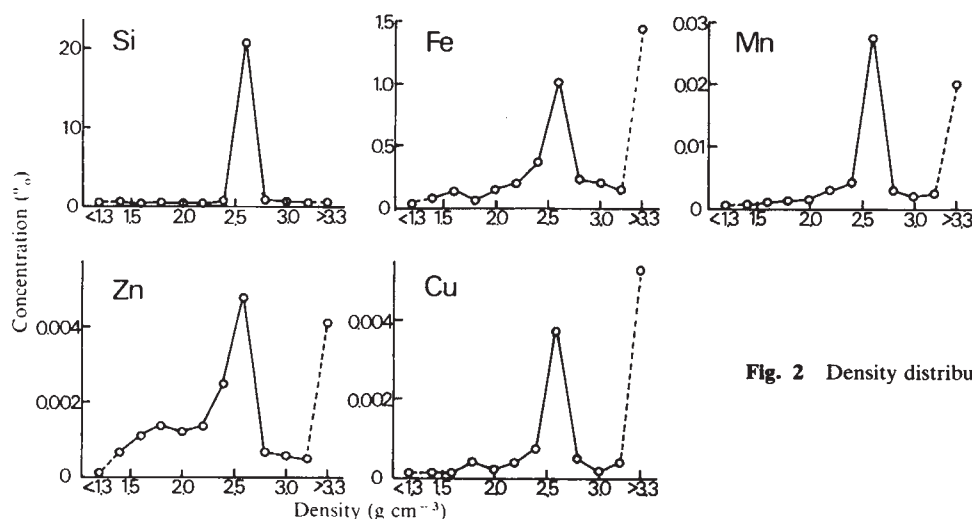


Fig. 2 Density distributions for Si, Fe, Mn, Zn and Cu in soil.

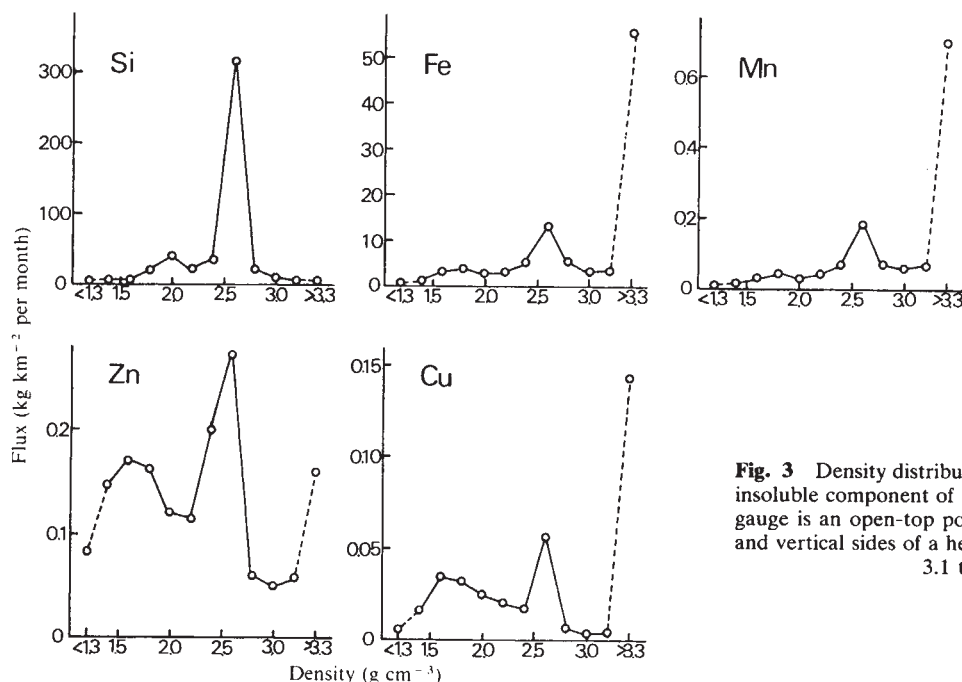


Fig. 3 Density distributions for Si, Fe, Mn, Zn and Cu in water-insoluble component of dustfall collected in a deposit gauge. The gauge is an open-top polyethylene cylinder of diameter 105 mm and vertical sides of a height of 215 mm. The flux of dustfall was 3.1 ton km^{-2} per month.

density higher than 3.3 g cm^{-3} . The density distribution of Mn particles in IPs favours low density. Manganese shows lower volatility than chalcophilic elements and is volatilized in a lesser extent in the high-temperature processes. However, low-density components are probably introduced into the atmosphere after being volatilized. The presence of dense particles ($>3.3 \text{ g cm}^{-3}$) in TSPs suggests that abrasional processes such as crustal weathering are also the sources of atmospheric Mn particles.

In urban area, soils, street dusts, construction dusts, and so on should be efficient sources of coarse particles, most of which will rapidly fall out near their sources, but some of which should suspend as atmospheric particulates. The density distributions of Si, Fe, Mn, Zn and Cu in soil samples are shown in Fig. 2. The maximum value was found for Si in the density range of $2.5\text{--}2.7 \text{ g cm}^{-3}$, corresponding to the density of SiO_2 . However, all elements with the exception of Si show two definite peaks in the density distributions at the density ranges of $2.5\text{--}2.7$ and $>3.3 \text{ g cm}^{-3}$. The presence of particles with density $>3.3 \text{ g cm}^{-3}$ would be explained by the presence of these elements as oxides. However, we are unaware of any crustal materials, containing these elements, with the density near 2.6 g cm^{-3} . Although the mechanisms responsible for the low density are not yet understood, the presence of these low-density particles would be explained by the presence of coarse SiO_2 particles which are coated with submicrometre platelets containing these elements. The particles associated with the density ranges of $2.5\text{--}2.7$ and $>3.3 \text{ g cm}^{-3}$ can be characterized by the natural sources such as wind-blown soil dusts. The soil particles floated into the atmosphere are likely to give additional peaks at the density ranges of $2.5\text{--}2.7$ and $>3.3 \text{ g cm}^{-3}$. The elemental density distributions of suspended particulates are completely different from those of soil. However, the density distributions of Fe and Mn in TSPs show a gradual slope in the density region around 2.6 g cm^{-3} , and imply the presence of soil-derived Fe and Mn particles.

For soil particles, gravitational fallout is an important removal mechanism. The elemental density distributions for the water-insoluble component of dustfall, illustrated in Fig. 3 are essentially similar to those for the soil. The matrix material of dustfall is considered to be composed primarily of soil-derived materials. However, the low-density fractions of dustfall seem to consist of carbonaceous materials judging from their black colour. The density distributions of chalcophilic elements Zn and Cu show a broad peak near the density of 1.6 g cm^{-3} , indicative of

anthropogenic sources, although the concentrations are still high in the density ranges of $2.5\text{--}2.7$ and $>3.3 \text{ g cm}^{-3}$.

Even though the most serious criticism of this technique is that the heavy liquids used for the density gradient separation would attack atmospheric particulates dissolving out some of the tarry matter and disagglomerating some clusters of particles, the technique is the only way to fractionate atmospheric particulates within the critical particle density range and provide valuable qualitative information on the physicochemical characteristics of atmospheric particulates. Until more quantitative techniques are developed, most quantitative tests for the formation and transport mechanisms for atmospheric particulates must rely on the concentration versus particle density data as well as on the concentration versus particle size data.

Received 27 June; accepted 7 November 1983.

- Sverdrup, G. M., Whitby, K. T. & Clark, W. E. *Atmos. Environ.* **9**, 483-494 (1975).
- Durham, J. L., Wilson, W. E., Ellestad, T. G., Willeke, K. & Whitby, K. T. *Atmos. Environ.* **9**, 717-722 (1975).
- Watson, J. G., Chow, J. C., Shah, J. J. & Pace, T. G. *J. Air Pollut. Control Ass.* **33**, 114-119 (1983).
- Jutze, G. A. & Foster, K. I. *J. Air Pollut. Control. Ass.* **17**, 17-25 (1967).
- Wedding, J. B., McFarland, A. R. & Cermak, J. E. *Envir. Sci. Technol.* **11**, 387-390 (1977).
- Dams, R., Rahn, K. A. & Winchester, J. W. *Envir. Sci. Technol.* **6**, 441-448 (1972).
- Olson, K. W. & Skogerboe, R. K. *Envir. Sci. Technol.* **9**, 227-230 (1975).
- Biggins, P. D. E. & Harrison, R. M. *Envir. Sci. Technol.* **13**, 558-565 (1979).
- Biggins, P. D. E. & Harrison, R. M. *Envir. Sci. Technol.* **14**, 336-339 (1980).
- Linton, R. W., Natusch, D. F. S., Solomon, R. L. & Evans, C. A. *Jr Envir. Sci. Technol.* **14**, 159-164 (1980).
- Hopke, P. K., Lamb, R. E. & Natusch, D. F. S. *Envir. Sci. Technol.* **14**, 164-172 (1980).
- Sugimae, A. *Taiki Osen Gakkaishi* **14**, 389-398 (1979).
- Sugimae, A. *ICP Inform. Newsl.* **6**, 619-644 (1981).
- Sugimae, A. *Taiki Osen Gakkaishi* **18**, 233-240 (1983).
- Sugimae, A. *Taiki Osen Gakkaishi* (in the press).
- Whytlaw-Gray, R. & Paterson, H. S. *Smoke* (Edward Arnold, London, 1932).
- Gladney, E. S., Zoller, W. H., Jones, A. G. & Gordon, G. E. *Envir. Sci. Technol.* **8**, 551-556 (1974).
- Dzubay, T. G. & Stevens, R. K. *Envir. Sci. Technol.* **9**, 663-668 (1975).
- Flocchini, R. G. *et al. Envir. Sci. Technol.* **10**, 76-82 (1976).
- Hardy, K. A., Akselsson, R., Nelson, J. W. & Winchester, J. W. *Envir. Sci. Technol.* **10**, 176-182 (1976).
- Paciga, J. J. & Jervis, R. E. *Envir. Sci. Technol.* **10**, 1124-1128 (1976).
- Kadowaki, S. *Envir. Sci. Technol.* **13**, 1130-1133 (1979).
- Tanaka, S., Darzi, M. & Winchester, J. W. *Atmos. Environ.* **14**, 1421-1426 (1980).
- Adams, F., Van Espan, P. & Maenhaut, W. *Atmos. Environ.* **17**, 1521-1536 (1983).
- Gillette, D. A., Blifford, I. H. & Fenster, C. R. *J. appl. Met.* **11**, 977-987 (1972).
- Rahn, K. A. *Atmos. Environ.* **10**, 587-601 (1976).
- Fisher, G. L., Chang, D. P. Y. & Brummer, M. *Science* **192**, 553-555 (1976).
- Campbell, J. A., Laul, J. C., Nielson, K. K. & Smith, R. D. *Analyt. Chem.* **50**, 1032-1040 (1978).
- Smith, R. D. *Prog. Energy Combust. Sci.* **6**, 53-119 (1980).

Photochemical production of carbonyl sulphide in marine surface waters

R. J. Ferek & M. O. Andreae

Department of Oceanography, Florida State University, Tallahassee, Florida 32306, USA

Carbonyl sulphide (COS) has been proposed as the major source of the stratospheric sulphate layer in times of low volcanic activity¹ because of its long tropospheric residence time², its uniform distribution in both the Northern and Southern Hemispheres³, and the lack of knowledge of any significant sinks except photodissociation in the stratosphere⁴. Known sources of COS include biomass burning⁵, volcanoes^{6,7}, salt marshes^{2,8}, fossil fuel combustion⁹, and oxidation of atmospheric CS₂ (refs 10, 11). Also, recent measurements¹²⁻¹⁴ have shown the surface oceans to be a source of COS. The total flux to the atmosphere from these known sources is substantially greater than the flux to the stratosphere needed to maintain the sulphate layer¹⁰. Thus questions remain about the sources and sinks of COS, and we have now performed a series of experiments in nearshore ocean and estuarine waters to investigate further the oceanic source of this gas. Dissolved COS in near-surface waters was found to be supersaturated with respect to atmospheric COS, and the dissolved concentrations varied diurnally with light intensity. Similar behaviour has been observed by other investigators for carbon monoxide, which is produced photochemically in surface seawater¹⁵. Furthermore, COS production appears to be independent of salinity, photo-synthetic activity and microbial activity, and is probably the result of photooxidation of organic matter. COS was produced in the laboratory by UV-irradiation of seawater and solutions of several biochemically relevant organosulphur compounds.

Measurements of both atmospheric and dissolved COS were performed using an instrumental system described in detail elsewhere¹⁴. For air measurements, a small volume (300 cm³) of ambient air was passed through a tube immersed in liquid nitrogen to trap out gases which were then separated on a chromatographic column and analysed by a flame photometric detector for sulphur. For dissolved COS, a closed volume of air was cycled through a gas bubbler in a glass column containing rapidly exchanging seawater until the air was equilibrated with the dissolved gases in the seawater. The equilibrated air was then analysed as before and compared with the ambient air concentration to obtain directly the COS saturation ratio (SR) in the water (that is $[\text{COS}]_{\text{equil air}}/[\text{COS}]_{\text{amb air}} = \text{SR}$). Liquid concentrations can be calculated from the saturation ratio using Henry's law as follows:

$$[\text{COS}]_{\text{liq}} = \frac{[\text{COS}]_{\text{amb air}}}{h} \times \text{SR}$$

where h is the Henry's law constant for COS in seawater¹⁶.

Initial measurements of the diurnal behaviour of dissolved COS were performed at the FSU Marine Laboratory located at Turkey Point on the north coast of the Gulf of Mexico south of Tallahassee, Florida. A continuous supply of seawater was obtained from the seawater circulation system with intake ~650 m offshore in 0.5–1.5 m of water. The residence time of water in the system was ~45 min. Figure 1 shows the results of two sampling periods at the Marine Laboratory. COS supersaturation shows a large diurnal cycle with peaks occurring during mid to late afternoon. No effect of tide (semidiurnal at this site) was observed.

Further measurements of COS saturation in surface waters were made during a cruise of the RV *Cape Hatteras* to Chesapeake Bay (28 March–3 April 1983) and Delaware Bay (4–8 April 1983). Seawater was obtained from the ship's continuously pumping seawater system with intake located 3 m below the surface. Ambient air was drawn through a Teflon

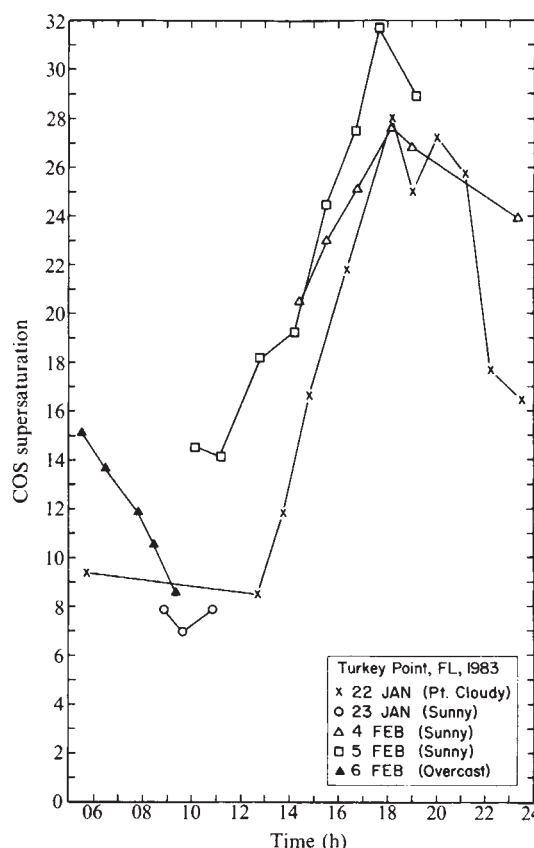


Fig. 1 Diurnal variation of COS in surface seawater measured at FSU Marine Laboratory, Turkey Point, Florida (northeastern Gulf of Mexico). The data are expressed as supersaturation ratio, that is, the ratio between the actual concentration in seawater and the concentration in equilibrium with ambient air.

tube (6.4 mm i.d.) which extended from the top of the mast to the laboratory. Figure 2 shows the diurnal behaviour of COS over a range of locations in both bays together with the hourly totals of light intensity in einsteins per m² (T. R. Fisher, personal communication). The highest COS supersaturations occur on days when light intensity was greatest. 28 and 29 March were both north-south transects of Chesapeake Bay over the same route. Although light measurements were not available on 28 March, it was a sunny day compared with partly cloudy conditions on 29 March, hence the higher COS levels on 28 March. Measurements on 30 March, 31 March, 1 April, and 2 April were taken at stations on Chesapeake Bay at surface salinities of 24, 17.4, 11, and 12‰, respectively. The magnitude of the COS saturation is again related to light intensity with no observable effect of salinity. Some of the variation on 1 April between 1100 and 1500 h may be due to vertical mixing as indicated by a lowering of the depth of the pycnocline between those hours. Since the secchi disk transparency was only slightly greater than 1 m, an increase in vertical mixing would be expected to mix in water from below with low COS concentration. This is reflected in the lowering of the COS saturation between 1100 and 1500 h on 1 April as opposed to a continuous rise during the daylight hours as illustrated in the FSU Marine Laboratory data (Fig. 1). The 3 April measurements were made near the mouth of the Susquehanna River (surface salinity = 0.2‰) on a sunny day. The COS saturation remained low relative to the other sunny days, probably due to low light penetration into the very turbid water (secchi disk transparency = 35 cm) at this station. 4 April was a north-south transect of Delaware Bay and 5 April was a station off Cape Henlopen (salinity = 29‰). COS saturations on both days do not appear appreciably different from the earlier Chesapeake stations.

Fig. 2 COS supersaturation and light intensity measured in surface waters during cruise of RV *Cape Hatteras* in Chesapeake and Delaware Bays, spring 1983.

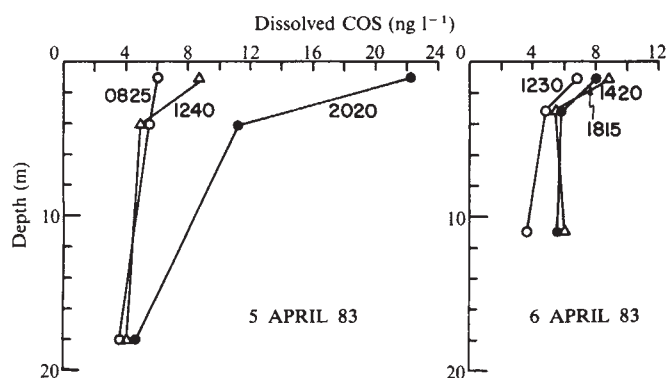
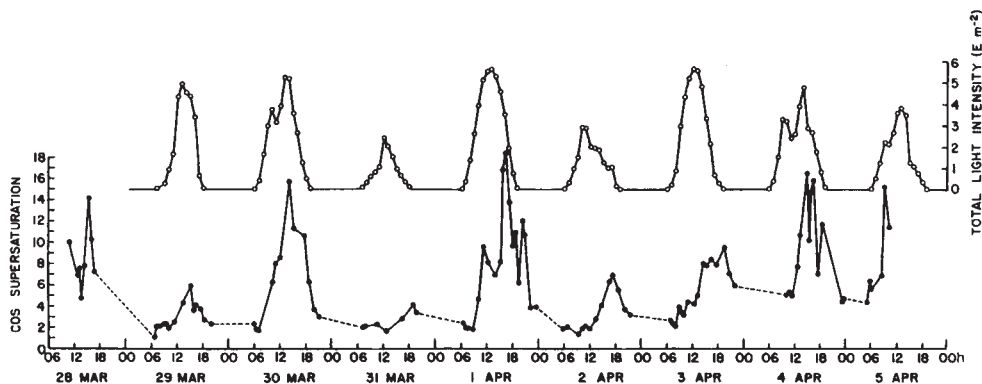
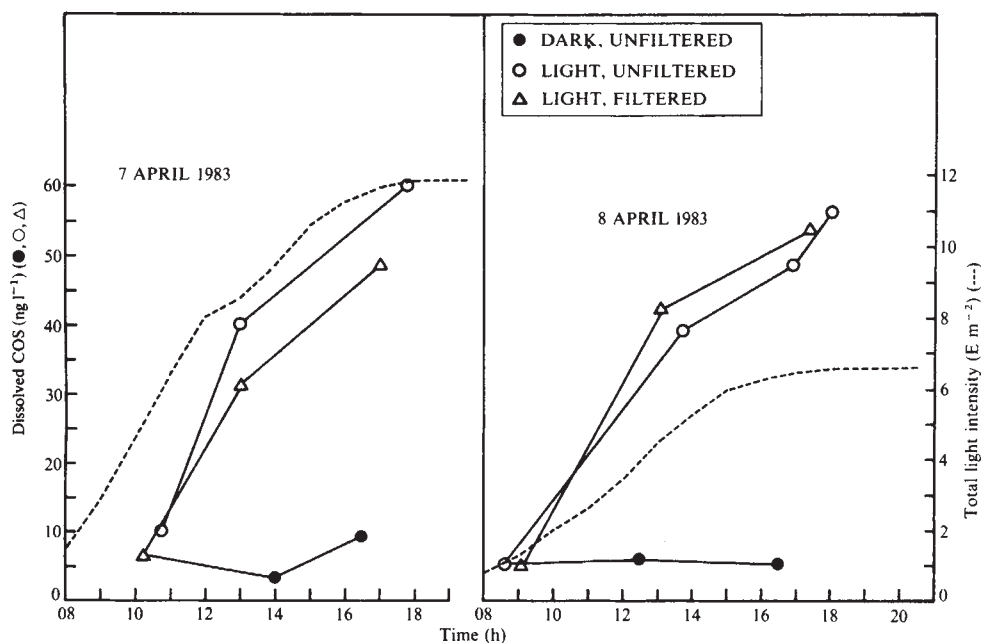


Fig. 3 Variation of the depth profile of dissolved COS concentration for two stations in lower Delaware Bay. 5 April was partly cloudy, 6 April was overcast. The time of sampling is indicated on each profile.

Fig. 4 Increase of COS concentration in surface seawater samples (filtered and unfiltered) incubated in light and dark bottles. The total integrated light intensity is indicated by a dotted line.



While at the Cape Henlopen station (5 April) and the 6 April station in lower Delaware Bay (surface salinity = 19‰), samples were collected at different depths by hydrocasts with 5-l Niskin bottles. 50-ml water samples were filtered through Gelman AE filters and purged with helium for 20 min in a gas stripping column with the effluent passed through a glass bead trap immersed in liquid nitrogen. Successive purges showed 20 min to be sufficient purging time to strip >90% of the dissolved COS from the samples. Figure 3 shows the results of these experiments. 5 April was a partly cloudy day and 6 April was overcast. Secchi disk transparencies at the two stations were 2.6 and 1.4 m, respectively. On both days COS at the deepest level

remained fairly constant and 2–5 times lower than the surface concentrations. Diurnal variation was greatest at the surface level, indicating that production of COS was confined to the photic zone.

A final experiment to check the effect of biological activity on the production of COS was performed on 7 and 8 April. Surface samples were collected at two stations in upper Delaware Bay (salinities 7.0 and 3.4‰). One sample was sterilized by filtering through a 0.45- μ m Nuclepore filter and, along with an unfiltered sample, analysed immediately by purging with He. Other filtered and unfiltered samples were placed in light and dark bottles and incubated in daylight for different lengths

of time. Figure 4 shows the results of these measurements. Also plotted are the running totals of light intensity throughout the day. The production of COS in the light bottles followed the light intensity, while dark bottles remained essentially constant at the initial COS concentration. There was essentially no difference between the filtered and unfiltered samples within the experimental uncertainty. The production of COS seems to be due to photooxidation of dissolved organic matter and thus only indirectly related to biological activity. As the 0.45- μm filter will remove all phytoplankton cells, but not all bacteria, a microbial production of COS cannot be rigorously excluded on the basis of this experiment. However, repeated irradiation of seawater with UV light from a small mercury discharge lamp ($\sim 400 \mu\text{W cm}^{-2}$ of 254 nm light) in the presence of dissolved O_2 , followed each time by degassing and determination of the COS produced, released COS in slowly decreasing amounts. Under the high UV intensities used, bacteria would have been killed during the first irradiation-degassing cycle and no further COS production should have taken place. UV irradiation of 100 μM solutions of several biochemically relevant organosulphur compounds (cysteine, methionine, glutathione, dimethylpropiothetin) in deionized, distilled water in the same conditions also produced substantial amounts of COS. This reaction did not occur when the solutions were stripped of dissolved O_2 by purging with He before irradiation.

Further evidence for the sea surface being a source of COS to the atmosphere can be seen in the ambient air concentrations measured at the three different locations. At Tallahassee, the mean ambient concentration of COS on 17 January, 18 January, and 18 February 1983 was $\bar{x} \pm s = 515 \pm 82$ p.p.t.v. (parts per 10^{12} by volume) ($n = 28$). This is very close to the mean tropospheric concentration of 512 ± 60 p.p.t.v. reported by Torres *et al.*³ for a total of 346 measurements during Project GAMETAG. During the period of sampling at the FSU Marine Laboratory, atmospheric concentrations were 537 ± 59 p.p.t.v. ($n = 65$), not significantly different from Tallahassee. However, during the Chesapeake and Delaware Bays cruise, the ambient air concentration was 627 ± 84 p.p.t.v. ($n = 204$), $\sim 20\%$ higher than the continental and coastal concentrations, and may be due to the closer proximity to the surface seawater source. In addition, air concentrations averaged 641 ± 88 p.p.t.v. ($n = 131$) between the hours of 10 am and 8 pm and 601 ± 70 p.p.t.v. ($n = 73$) between 8 pm and 10 am.

Earlier measurements^{12,14} in open ocean areas found COS saturation in surface waters averaging approximately 2.0, which results in a flux to the atmosphere of $\sim 0.5\text{--}0.9 \times 10^{12}$ g COS yr^{-1} . The higher saturations observed in coastal waters may make the total flux from the oceans greater than previously estimated. We are attempting to gather data from more diverse areas of the ocean to further refine our estimate of this flux.

We thank the captain and crew of the RV *Cape Hatteras* and chief scientist T. R. Fisher for their cooperation. J. T. Byrd and G. Baptist helped with sample collection. This work was supported by NSF grant ATM-8017574.

Received 14 July; accepted 15 September 1983.

- Crutzen, P. J. *Geophys. Res. Lett.* **3**, 73–76 (1976).
- Aneja, V. P., Aneja, A. P. & Adams, D. F. *J. Air. Pollut. Control Ass.* **32**, 803–806 (1982).
- Torres, A. L., Maroulis, P. J., Goldberg, A. B. & Bandy, A. R. *J. geophys. Res.* **85**, 7357–7360 (1980).
- Turco, R. P., Whitten, R. C., Toon, O. B., Pollack, J. B. & Hamill, P. *Nature* **283**, 283–286 (1980).
- Crutzen, P. J., Heidt, L. E., Krasnec, J. P., Pollock, W. H. & Seiler, W. *Nature* **282**, 253–256 (1979).
- Cadle, R. D. *Rev. Geophys. Space Phys.* **18**, 746–752 (1980).
- Rasmussen, R. A., Khalil, M. A. K., Dalluge, R. W. & Penkett, S. A. *Science* **215**, 665–667 (1982).
- Carroll, M. A., Heidt, L. E., Cicerone, R. J. & Prinn, R. G. *EOS* **63**, 893 (1982).
- Turco, R. P., Cicerone, R. J., Inn, E. C. Y. & Capone, L. A. *J. geophys. Res.* **86**, 5373–5377 (1981).
- Sze, N. D. & Ko, M. K. W. *Nature* **280**, 308–310 (1979).
- Logan, J. A., McElroy, M. B., Wofsy, S. C. & Prather, M. J. *Nature* **281**, 185–188 (1979).
- Rasmussen, R. A., Khalil, M. A. K. & Hoyt, S. D. *Atmos. Envir.* **16**, 1591–1594 (1982).
- Johnson, J. E., Harrison, H. & Heidt, L. E. *EOS* **63**, 894 (1982).
- Ferek, R. J. & Andreae, M. O. *Geophys. Res. Lett.* **10**, 393–396 (1983).
- Conrad, R., Seiler, W., Bunse, G. & Giehl, H. *J. geophys. Res.* **87**, 8839–8852 (1982).
- Rasmussen, R. A., Hoyt, S. D. & Khalil, M. A. K. *Chemosphere* **11**, 869–875 (1982).

First observations of the interaction of ocean swell with sea ice using satellite radar altimeter data

C. G. Rapley*

Groupe de Recherche de Géodésie Spatiale,
18 Avenue Edouard Belin, 31055 Toulouse Cedex, France

The action of ocean waves and swell on sea ice in the marginal ice zone (MIZ) controls the size distribution of ice floes and thus affects ice dynamics and the thermodynamics of ice growth and decay. An understanding of ice processes is important because of the influence of the sea-ice cover on high latitude weather and global climate¹. Also, long-term changes in average sea ice extent could provide a sensitive indication of climatic change if the degree of shorter term variability were better understood². Previous observations of the propagation of swell within the ice pack have been limited in spatial and temporal extent^{3–6}. Here we present new results from the US Seasat satellite which suggest that radar altimetry can provide a powerful means of global synoptic monitoring of the interaction between ocean and ice.

The US satellites GEOS-3 and Seasat carried pulse-limited altimeters designed for studies of the geoid and ocean surface⁷. Both instruments obtained data over regions of sea ice. The radar echoes observed were highly variable but were generally stronger and more peaked in time than those from open water⁸. Brown⁹ attributes the intense component of first-year ice returns to a coherent signal from smooth ice platelets all lying at the same level amongst the floes. Robin *et al.*¹⁰ show that the peaked component will dominate over the diffuse echo if even as little as 0.01% of the surface contributes a coherent signal. Thus, provided any small areas of new ice or calm water are exposed within the altimeter footprint, we may expect to observe a strong return, and if the sea surface between floes is flat the waveform will be highly peaked.

Figure 1 shows three examples of the evolution of on-board computed values of 'significant wave height' (SWH) and signal strength (automatic gain control, AGC) as the Seasat altimeter passed from open ocean into the ice pack. Obvious features are the step in signal strength at the ice edge (indicated by the arrows) and the greater variability of both parameters over the ice. A particularly interesting feature of the SWH data in Fig. 1a and b is the gradual overall decline to zero with increasing distance from the ice edge. It is not immediately obvious to what property of the pack ice the SWH value corresponds. An insight into the effect is provided by the altimeter waveform data which typically evolve as shown schematically in Fig. 2 (see also Fig. 1d, e of Robin *et al.*¹⁰). Pulse shapes just inside the ice edge (Fig. 2a) are often almost identical to those from the nearby ocean, although amplitudes may be as much as 10 times greater. Altimeter pulse risetimes are a measure of the height probability density function of reflecting elements at nadir, and pulse fall times depend on the slope distribution (that is, for specular reflection the maximum duration of the pulse is given by $t_m = (h/c)\theta_{\max}^2$ where $\theta_{\max}$ is the maximum surface slope, and c and h are the velocity of light and satellite altitude respectively). Apparently, neither distribution changes significantly across the ice edge, although the large increase in signal strength implies that the nature of the reflecting elements must alter dramatically. This may be naturally explained if most of the power returned is reflected from thin ice platelets or smooth water between floes as discussed previously. The increasing leading edge slope and more rapid decay of waveforms deeper in the pack (Fig. 2b,c) then imply a continuous reduction in the interrelated height and slope distributions of the ocean

* Permanent address: Mullard Space Science Laboratory, Department of Physics and Astronomy, University College London, Dorking, Surrey RH5 6NT, UK.

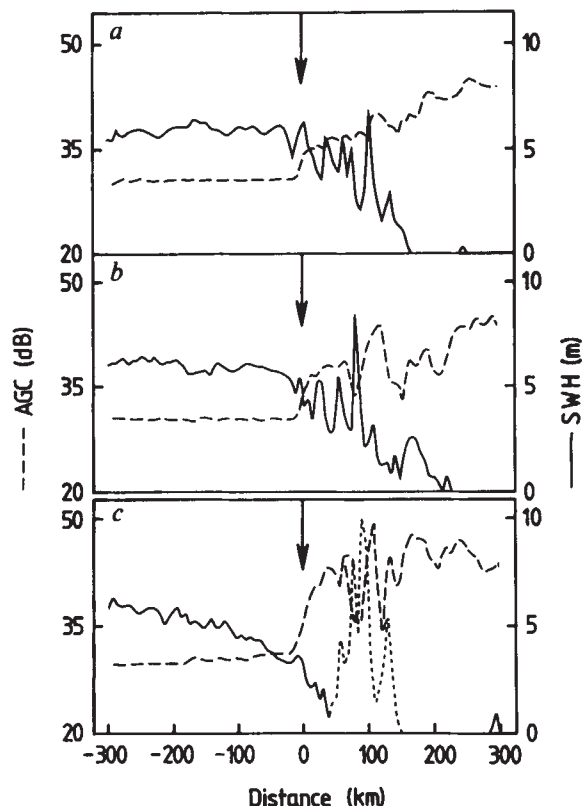


Fig. 1 Three examples of the evolution of on-board computed SWH and AGC values as the Seasat altimeter passed from ocean to sea ice. The ice edge is indicated by the arrow in each case. Note the sudden increase in AGC at the edge and the gradual overall decline of SWH value to zero. Note also the increased SWH noise within the ice pack and the poor data points, particularly in example c.

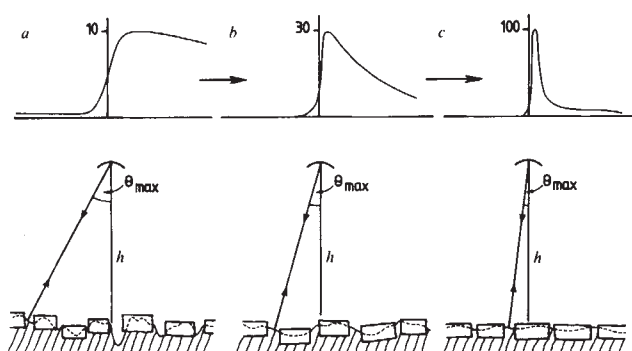


Fig. 2 The evolution of waveform shape and strength as a function of increasing distance into the ice pack. Leading-edge slopes depend on the height distribution of scatterers at nadir just as for ocean returns. Falltimes are determined by the distribution of surface slopes as shown schematically in the lower panels. Pulse strengths are relative to a value unity for a standard ocean return. Pulse shape and strength change as ocean waves and swell are damped out by the action of ice floes within the pack.

(or thin ice) surface as a function of distance from the ice edge. This is consistent with the known behaviour of ocean waves within the ice^{5,6}.

The relationship between pulse shape and SWH value is not straightforward. The on-board processing methods used by the Seasat altimeter were based on a standard ocean waveform model and the measured instrument performance characteristics^{11,12}. An analysis of the response to non-standard pulses is complicated by interactions between all three processing loops¹³. However, an SWH output of zero was special and indicated a value less than the minimum resolvable by the instrument (~ 40 cm). This could only occur for pulses more peaked than

an ocean return, although the degree of 'peakiness' necessary was a function of pulse risetime. Nevertheless, risetime and peakiness are partially interrelated as we have described, and we may therefore use zero SWH as a crude reference level of surface flatness.

Unfortunately, the on-board computed SWH values over sea ice suffered from increased noise and poor data points (note Fig. 1c in particular). The latter consist of correlated excursions of the height and SWH signals and resulted from brief attempts by the on-board tracker to follow receding bright features during transitions between areas of high and low surface reflectivity. The effect is similar to that described by Thomas *et al.*¹⁴ for transitions across an ice front. The geographical distribution of extreme events shows a strong concentration in the vicinity of the ice edge and in areas known to be prone to polynya formation, implying an association with transitions from new or first-year ice to open water. Alternatively, 'hunting' by the on-board tracker when processing highly peaked pulses¹³ could have introduced similar errors. Occurrences within the ice pack were more frequent towards the coast where AGC values were low, and may have resulted from transitions to areas of continuous ice or snow cover. Much of the contaminated data can be recovered by analysis of the altimeter waveforms but this represents a very major computing task. Thus, an exploratory investigation of swell propagation within the ice was carried out using the SWH data, recognizing its limitations. The use of a spatial resolution of ~ 70 km (10 point averages) and the averaging effect of adjacent groundtracks reduced the impact of data contamination on the results.

Figure 3 shows data from four successive repeats of the 3-day orbit cycle executed by Seasat during September 1978 at the time of maximum sea ice extent. The ice edge derived using an AGC threshold of 33 dB and the zero SWH contour are shown, the region between being shown as solid areas. Despite the coarse spatial resolution used, inconsistencies in the location of the ice edge occur due to motion within a 3-day cycle. These are represented by small re-entrant features. The SWH contour joins locations at which zero SWH is first encountered on entry into the ice pack. All non-zero SWH values within this contour are ignored as are occasional groundtracks obviously affected by poor data. Also shown are zones of high ocean swell (> 5 m) derived using the method of Mognard¹⁵ (see also ref. 16). The zones shown represent major storm systems. An immediate qualitative correspondence can be seen between areas in which the zero SWH contour is furthest from the ice edge and nearby zones of high swell. Furthermore, changes in the width of the band follow the temporal and spatial evolution of the swell storms. For example, the sector between 30° E and 60° W contains little swell storm activity and consistently exhibits a narrow band of swell penetration, whereas the decay of the major swell storm between 30° E and 120° E is matched by a decline in penetration distance.

The altimeter data therefore provide a means of using the ice as a detector of swell or, alternatively, the swell as a probe of the ice. The energy decay of each swell component is known to be exponential with an attenuation coefficient strongly dependent on swell wavelength and ice thickness^{5,6}. The width of the band shown here also depends on the amplitude of the penetrating component at the ice edge and may be affected by the inadequacies of the SWH data as described above. Nevertheless, the observed penetration distances of up to 300–400 km are consistent with previous *in situ* measurements³. Detailed, quantitative studies will require analysis of the waveform data. Even so, the SWH data provide useful qualitative insights into ice properties. For example, the strong propagation of low swells into the edge features seen between 30° W and 90° W in Fig. 3a and at 30° W in Fig. 3d, implies that the ice is thin, as does their rapid spatial variability.

In an attempt to relate swell effects to ice edge growth, Fig. 4 shows a detail of ice edge changes over the four periods. A good general correlation is observed between ice edge variability and regions affected by swell, particularly in the case of the

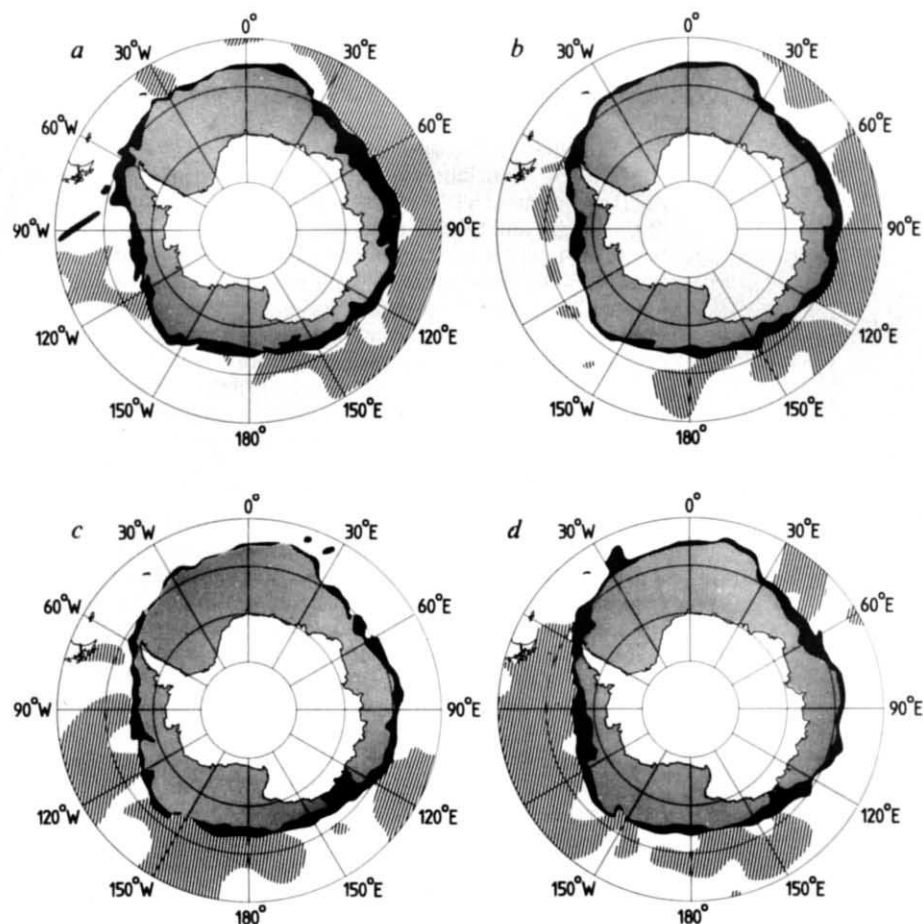


Fig. 3 Four consecutive 3-day maps of Antarctic sea ice (shaded areas) and regions of ocean swell greater than 5 m (cross-hatched areas) during September 1978. Zones of swell penetration into the ice (solid areas) derived from the SWH data contract and expand with the decline and subsequent growth of nearby ocean swell storms. Note the generally limited penetration in the region 30°E to 60°W where swell activity is low during this period, other than the feature suspected to consist of thin ice in *d*.

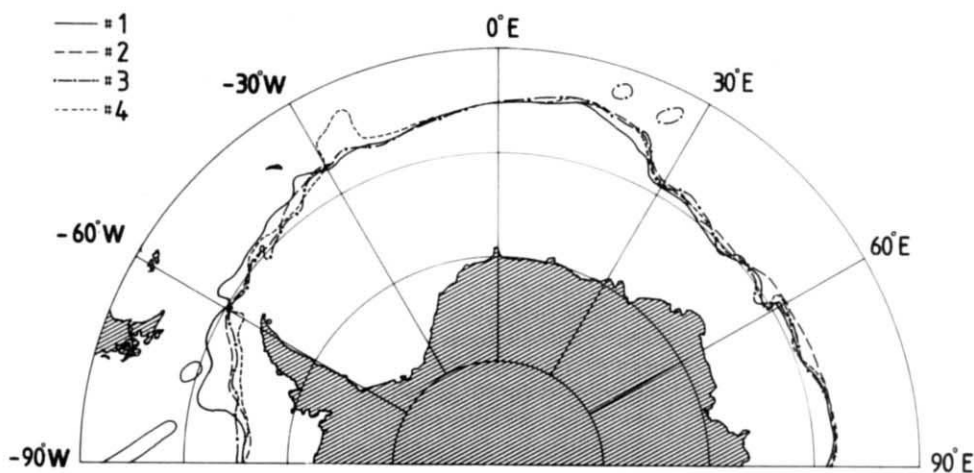


Fig. 4 Detail of the ice-edge variations over the four 3-day cycles. Greater variability occurs in regions affected by swell. Features showing the greatest changes are deeply penetrated by low-amplitude swells (see Fig. 3) implying that the ice is comparatively thin.

(apparently) thin ice features already noted. Keliher¹⁷ has previously reported an association between pack ice disturbances and regions of high sea state.

We may conclude, therefore, that variations in altimeter return pulse shape provide a powerful means of monitoring swell propagation within the ice pack. Difficulties associated with the use of on-board SWH values have important implications on instrument design and data analysis procedures for

future satellite altimeter missions.

I thank the Centre Nationale d'Etudes Spatiales, France for financial support and the Groupe de Recherche de Géodésie Spatiale for their hospitality and assistance. I am particularly grateful to Mr Claude Brossier for his hard work and enthusiasm in producing the data plots and to Dr Michel Lefebvre for his encouragement. I acknowledge useful discussions with Drs Nelly Mognard and Bill Campbell.

Received 19 September; accepted 28 October 1983.

1. Committee on Glaciology, Polar Research Board Snow and Ice Research, An Assessment (National Academy Press, Washington DC, 1983).
2. Zwally, H. J., Parkinson, C. L. & Comiso, J. C. *Science* **220**, 1005-1012 (1983).
3. Robin, G. de Q. *Phil. Trans. R. Soc. A* **255**, 313-339 (1963).
4. Wadhams, P. *J. geophys. Res.* **80**, 4520-4528 (1975).
5. Wadhams, P. *Deep Sea Res.* **25**, 23-40 (1978).
6. Squire, V. A. & Moore, S. C. *Nature* **283**, 365-368 (1980).
7. Townsend, W. F. *IEEE J. Oceanic Engng* **OE-5**(2), 80-92 (1980).
8. Dwyer, R. E. & Godin, R. H. *NASA Contract Report* 156862 (1980).

9. Brown, G. S. *Radio Sci.* **17**, 233-243 (1982).
10. Robin, G. de Q., Drewry, D. J. & Squire, V. A. *Phil. Trans. R. Soc. A* **309**, 447-463 (1983).
11. MacArthur, J. L. *Proc. Oceans-76, MTS-IEEE* 1-8 (1976).
12. MacArthur, J. L. *Applied Physics Laboratory, SDO-5232* (Johns Hopkins University, 1978).
13. Rapley, C. G. *et al. ESA Contract Report* 5182/82/F/CG(SC) (1983).
14. Thomas, R. H., Martin, T. V. & Zwally, J. H. *Ann. Glaciol.* (in the press).
15. Mognard, N. thesis, Univ. Paul Sabatier de Toulouse (1982).
16. Mognard, N., Campbell, W. J., Cheny, R. E. & Marsh, J. G. *J. geophys. Res.* **88**, 1736-1744 (1983).
17. Keliher, T. E. *AIDJEX Bull.* **34**, 114-135 (1976).

Earthquake seasonality before the 1906 San Francisco earthquake

Patrick H. McClellan

US Department of Interior, Minerals Management Service,
345 Middlefield Road, Menlo Park, California 94025, USA

Since 1855, 40 earthquakes of magnitude ≥ 5.5 have occurred on the San Andreas Fault (SAF) system in coastal central California, between latitudes $36^{\circ}30' \text{ N}$ and $39^{\circ}30' \text{ N}$ (Table 1)¹. Of the included 27 'independent' earthquakes, which occurred more than 1 yr after or 1 geocentric degree distant from a previous ≥ 5.5 mag earthquake, 3 occurred in summer, 7 in autumn, 4 in winter and 13 in spring. All of the spring earthquakes occurred during the 25-yr interval that culminated with the great San Francisco earthquake (8.25 mag) in the spring of 1906 (Fig. 1a, b). I report here that the observed frequency of the spring earthquakes before 1906 significantly exceeds (binomial $P=0.002$) seasonal frequencies expected from random variation in the earthquake rate, and thus invites serious inquiry into the subject of earthquake seasonality. More importantly, the time-clustering of spring earthquakes before 1906 raises the question of whether accelerated seasonal seismicity may herald the end of the 'seismic cycle' suspected¹ to modulate the long-term pattern of seismicity on the northern segment of the SAF system.

To avoid the *a priori* assumptions inherent in seasonal grouping of the data, the significance of the apparent earthquake seasonality before 1906 was further assessed with the Schuster test², which, as used here, measures the tendency of earthquakes to occur near the same time ('phase') of the year. Each earthquake datum within a year was converted to a phase angle ω_i (1 January = 0°) and the n phases were vectorially added. If

$$A = \sum_{i=1}^n \cos \omega_i \quad \text{and} \quad B = \sum_{i=1}^n \sin \omega_i$$

then the resultant vector has a net phase $\phi = \arctan(B/A)$ and a length $R = \sqrt{A^2 + B^2}$. The probability of observing a vector longer than R (that is, a better correlation) from a set of n random phases is $P = \exp(-R^2/n)$, where $n \geq 10$ (refs 3–5). The results are listed in Table 2. For the complete set of $n=27$ independent earthquakes since 1855, $\phi = 119^{\circ}$ (corresponding to a date of 30 April) with a random probability $P=0.21$ (not significant). However, for the 19 earthquakes between 1855 and the end of 1905, $\phi = 104^{\circ}$ and $P=0.029$. In other words, leading into the year 1906, there was a significant tendency for moderate and large earthquakes to occur not only in the spring but around the date of 15 April. The great San Francisco earthquake occurred on 18 April 1906.

Considering only the 'strongest' independent earthquakes (≥ 6.0 mag, as defined, for example, in ref. 6) between 1855 and 1906 ($n=7$) leaves too few events ($n < 10$) to be tested reliably for correlation by the Schuster method⁵. If the magnitude threshold is lowered below 6.0 mag, to enlarge the set of strongest events between those years, seasonality becomes increasingly apparent. Although clustering is not evident for the 10 earthquakes of ≥ 5.9 mag ($P=0.35$), it is significant at the $P=0.07$ level for the 13 events of ≥ 5.7 mag, and, as noted above, at the $P=0.029$ level for the 19 events of ≥ 5.5 mag. This indicates that the weaker earthquakes were, in large part, responsible for the observed seasonality. Alternatively, the set of seven strongest (≥ 6.0 mag) earthquakes before 1906 may be enlarged by adding to it the three strongest known pre-1855 central California earthquakes (Table 3), which reasonably are assumed also to have exceeded 6.0 mag⁷. Those 10 strongest earthquakes collectively are correlated, albeit marginally ($P=0.07$ – 0.09 ; Table 2), about the dates of 23–28 May, with most

Table 1 Earthquakes (≥ 5.5 mag) in coastal central California from $36^{\circ}30' \text{ N}$ to $39^{\circ}30' \text{ N}$, 1855–1982

Year	Date	Time	Latitude	Longitude	Magnitude
1856*	2 15	13:25	$37^{\circ}36' \text{ N}$	$122^{\circ}24' \text{ W}$	5.9
1858*	11 26	8:35	$37^{\circ}36' \text{ N}$	$122^{\circ}00' \text{ W}$	5.9
1864*	2 26	8:40	$36^{\circ}30' \text{ N}$	$121^{\circ}30' \text{ W}$	5.8
1865*	10 8	20:46	$37^{\circ}06' \text{ N}$	$121^{\circ}54' \text{ W}$	6.2
1866	3 26	20:12	$37^{\circ}06' \text{ N}$	$121^{\circ}30' \text{ W}$	5.8
1866	7 15	6:30	$37^{\circ}18' \text{ N}$	$121^{\circ}12' \text{ W}$	5.7
1868*	10 21	15:53	$37^{\circ}42' \text{ N}$	$122^{\circ}06' \text{ W}$	6.7
1870*	2 17	20:12	$37^{\circ}12' \text{ N}$	$122^{\circ}00' \text{ W}$	5.5
1881*	4 10	10:00	$37^{\circ}18' \text{ N}$	$121^{\circ}18' \text{ W}$	6.0
1882	3 6	21:45	$36^{\circ}42' \text{ N}$	$121^{\circ}12' \text{ W}$	5.5
1883*	3 30	15:45	$36^{\circ}54' \text{ N}$	$121^{\circ}36' \text{ W}$	5.7
1885*	3 31	7:56	$36^{\circ}42' \text{ N}$	$121^{\circ}18' \text{ W}$	5.6
1885	4 2	15:25	$37^{\circ}00' \text{ N}$	$121^{\circ}24' \text{ W}$	5.5
1889*	5 19	11:10	$38^{\circ}00' \text{ N}$	$121^{\circ}54' \text{ W}$	5.6
1890*	4 24	11:36	$36^{\circ}54' \text{ N}$	$121^{\circ}36' \text{ W}$	5.9
1891*	10 12	6:28	$38^{\circ}18' \text{ N}$	$122^{\circ}18' \text{ W}$	5.6
1892	4 19	10:50	$38^{\circ}30' \text{ N}$	$122^{\circ}00' \text{ W}$	6.8
1892	4 21	17:43	$38^{\circ}36' \text{ N}$	$121^{\circ}54' \text{ W}$	6.3
1897*	6 20	20:14	$36^{\circ}54' \text{ N}$	$121^{\circ}24' \text{ W}$	6.0
1898*	3 31	7:43	$38^{\circ}12' \text{ N}$	$122^{\circ}24' \text{ W}$	6.3
1898*	4 15	7:07	$39^{\circ}12' \text{ N}$	$123^{\circ}48' \text{ W}$	6.7
1899*	4 30	22:41	$36^{\circ}54' \text{ N}$	$121^{\circ}36' \text{ W}$	5.7
1899*	6 2	7:19	$37^{\circ}48' \text{ N}$	$122^{\circ}36' \text{ W}$	5.5
1899	7 6	20:10	$36^{\circ}42' \text{ N}$	$121^{\circ}18' \text{ W}$	5.5
1902*	5 19	18:31	$38^{\circ}18' \text{ N}$	$121^{\circ}54' \text{ W}$	5.5
1903*	6 11	13:12	$37^{\circ}36' \text{ N}$	$121^{\circ}48' \text{ W}$	5.6
1903	8 3	6:49	$37^{\circ}18' \text{ N}$	$121^{\circ}48' \text{ W}$	5.6
1906*	4 18	13:12	$37^{\circ}42' \text{ N}$	$122^{\circ}30' \text{ W}$	8.3
1910*	3 11	6:52	$36^{\circ}54' \text{ N}$	$121^{\circ}48' \text{ W}$	5.7
1911*	7 1	22:00	$37^{\circ}15' \text{ N}$	$121^{\circ}45' \text{ W}$	6.0
1914*	11 9	2:31	$37^{\circ}10' \text{ N}$	$122^{\circ}00' \text{ W}$	5.5
1926*	10 22	12:35	$36^{\circ}37' \text{ N}$	$122^{\circ}20' \text{ W}$	6.2
1926	10 22	13:35	$36^{\circ}34' \text{ N}$	$122^{\circ}20' \text{ W}$	6.2
1926	10 24	22:51	$37^{\circ}01' \text{ N}$	$122^{\circ}12' \text{ W}$	5.6
1927	2 15	23:54	$36^{\circ}57' \text{ N}$	$122^{\circ}15' \text{ W}$	5.6
1955*	9 5	2:01	$37^{\circ}22' \text{ N}$	$121^{\circ}47' \text{ W}$	5.6
1969*	10 2	4:56	$38^{\circ}28' \text{ N}$	$122^{\circ}41' \text{ W}$	5.5
1979*	8 6	17:05	$37^{\circ}05' \text{ N}$	$121^{\circ}28' \text{ W}$	5.9
1980	1 24	19:00	$37^{\circ}50' \text{ N}$	$121^{\circ}48' \text{ W}$	5.8
1980	1 27	2:33	$37^{\circ}51' \text{ N}$	$121^{\circ}47' \text{ W}$	5.5

Data are from the catalogue in ref. 1, Appendix, Table A1. All magnitudes before 1900 and some during 1900–27, are intensity magnitudes (M_I) estimated from published felt-area data on the basis of empirical relations between local magnitude (M_L) and extent of isoseismal areas for twentieth-century earthquakes in California and western Nevada¹⁴. A test of M_I against M_L for twentieth-century earthquakes in coastal central California¹ suggests that, for this subregion, M_I is slightly conservative and may underestimate M_L by about 0.2 ± 0.2 units. Thus, the intensity magnitudes listed are presumed to be not greatly biased relative to the instrumental magnitudes. Magnitude listed for the 1906 event is surface-wave magnitude (M_S), M_L recently determined at 6.9 (ref. 15).

* 'Independent' events, which were separated by more than 1 yr or 1 geocentric degree.

occurring in the spring season within the quarter-century before 1906.

Possible explanations for the observed earthquake seasonality include (1) incompleteness of the ≥ 5.5 mag earthquake catalogue, (2) fortuitous sampling of the catalogue by the assumed criteria of event independence, (3) fortuitous sampling of a limited geographical segment of the SAF system, (4) a real, though improbable, random variation in the earthquake occurrence rate and (5) a fundamental geological process. Because the catalogue is believed¹ to be complete for events down to 5.5 mag, it is unlikely that the observed high spring seismicity is due to the omission of numerous large earthquakes in other seasons. Likewise, the pattern probably is not an artefact of fortuitous sampling of the ≥ 5.5 mag catalogue, inasmuch as assuming the least conservative definition of event independence (Fig. 1c) also results in a disproportionately high frequency of

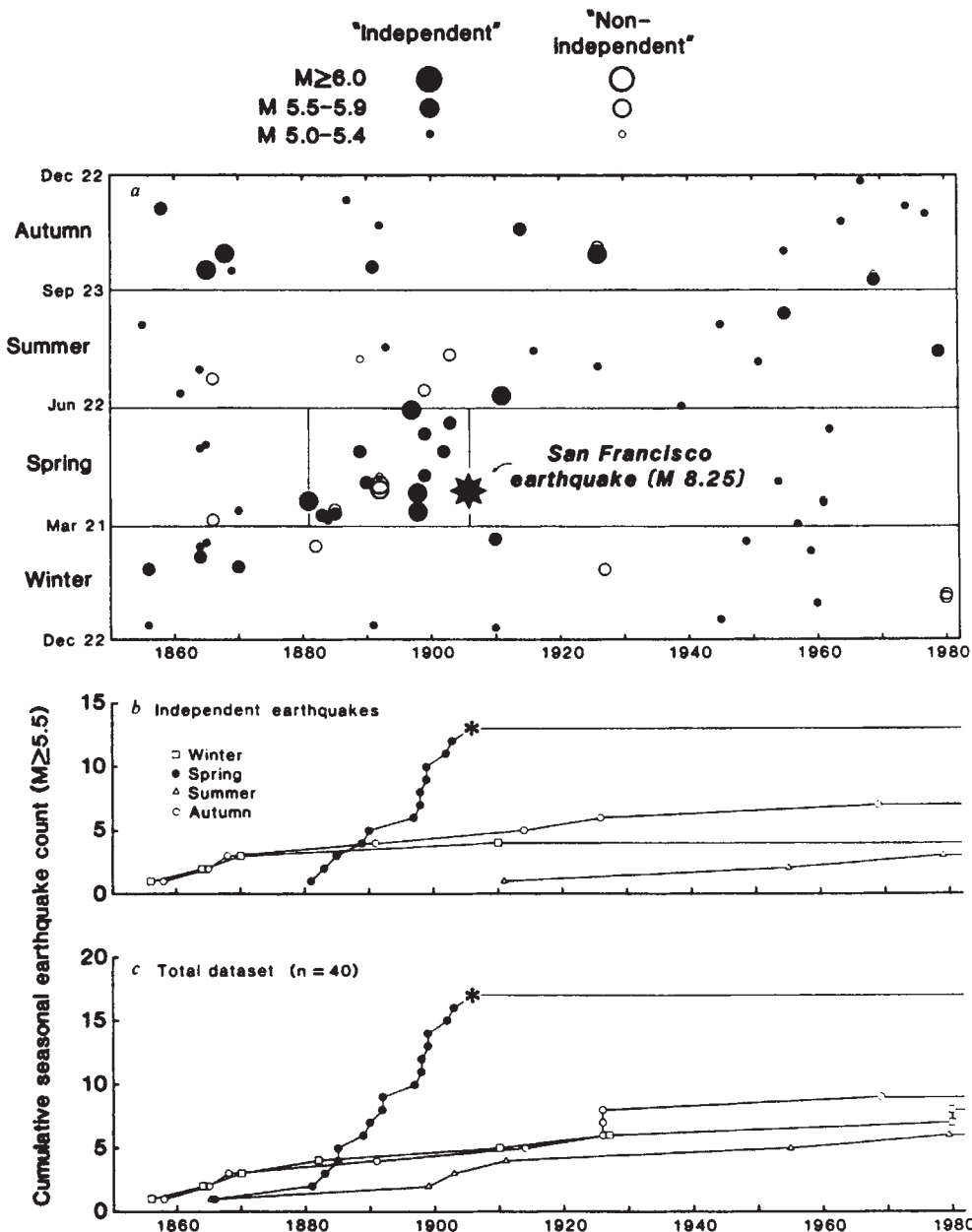


Fig. 1 *a*, Year-season plot of moderate to large (≥ 5.5 mag) earthquakes in coastal central California ($36^{\circ}30'-39^{\circ}30' N$) from 1855 to 1983. Independent ≥ 5.5 mag events are defined as those separated by more than 1 yr or 1 geocentric degree. Seasons are defined by the conventional dates of the winter solstice (22 December), vernal equinox (21 March), summer solstice (22 June) and autumnal equinox (23 September). *b*, Cumulative count by season of independent ≥ 5.5 mag earthquakes plotted in *a* ($n = 27$). *c*, Cumulative count by season of total ≥ 5.5 mag earthquake set plotted in *a* ($n = 40$).

spring earthquakes. A more conservative assumption than was used seems unreasonable.

Regarding the geographical limitation of the earthquake sample, Schuster analysis of independent strong earthquakes (≥ 6.0 mag) on the southern SAF system ($33-36^{\circ} N$, $n = 13$)⁶, beginning with the great 1857 earthquake there to the present time, show that those events also are correlated ($P = 0.04$) about a date in the spring (31 March), as are those earthquakes in combination with the 10 strong earthquakes since 1855 in central California ($36^{\circ}30'-39^{\circ}30' N$; Table 1) plus 12 other post-1855 events of about ≥ 6.0 mag which occurred on the SAF system between $33-36.5^{\circ} N$ (Table 3)⁷, but were not included in the list in ref. 6. This total set of ≥ 6.0 mag earthquakes ($n = 35$) is correlated ($P = 0.04$) about the date of 20 April, and suggests that the observed earthquake seasonality in coastal central California is stable against geographical extensions of the earthquake sample, at least for the strongest historical events over the extent of the SAF system. (Smaller earthquakes throughout the system were not analysed.) Finally, although the 25-yr-long sequence of spring earthquakes in central California may have been a mere coincidence of random events, the termination of the sequence with the unique 1906 earthquake circumstantially suggests otherwise. Thus, coastal central California appears to have been tectonically predisposed to the frequent

occurrence of moderate to large earthquakes during the spring season from 1881 to 1906.

A physical basis for the observed earthquake seasonality on the northern segment of the SAF system can only be conjectured at present. Nevertheless, the fundamental mechanism almost certainly involves the manner in which elastic strain accumulates in the crust along that portion of the North American-Pacific plate boundary. Crustal strain accumulating there is released primarily in ground-rupturing earthquakes on the SAF that are probably similar in size to the 1906 earthquake and are believed to occur there at intervals averaging from 150 yr^1 to perhaps 230 yr^8 in length. This range of estimated recurrence intervals approximates the average of 145 yr (range 100–225 yr) between large or great ground-rupturing earthquakes on the south-central SAF, like the 1857 event, during the past 17 centuries (K. E. Sieh, personal communication). In addition, in coastal central California, overall historical seismicity above the 5.0 mag level was generally higher during the half-century before the 1906 earthquake than in the half-century afterward¹, not unlike a pattern, called the 'seismic cycle'^{9,10}, observed before and after great earthquakes in the Wadati-Benioff zones of the western Pacific basin. However, an unequivocal precursory increase in seismicity, as characterizes 'Stage II' of this cycle¹⁰, has not hitherto been identified in pre-1906 earthquake data

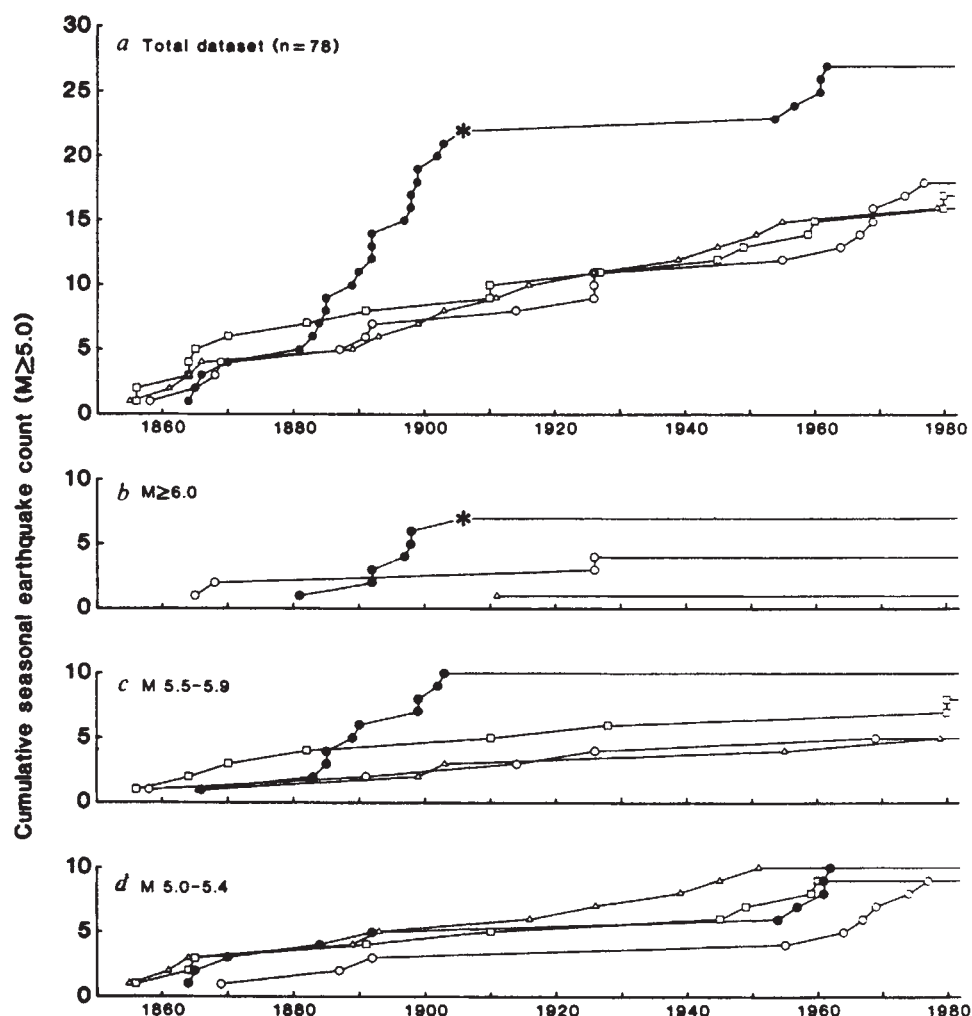


Fig. 2 *a*, Cumulative count by season of ≥ 5.0 mag earthquakes in coastal central California from 1855 to 1983, listed in ref. 1 ($n=78$). *b*, Total ≥ 6.0 mag earthquake subset in *a*. *c*, Total 5.5–5.9 mag earthquake subset in *a*. *d*, Total 5.0–5.4 mag earthquake subset in *a*.

from central California^{1,11}. Nevertheless, the quasiperiodic recurrence of great earthquakes on the northern SAF is suspected¹ to control a seismic cycle there also.

Conjectural mechanisms which might have selectively triggered earthquakes in the spring on faults in the SAF system necessarily presuppose that, before triggering, those faults were already tectonically strained to near the point of shear failure. It is not unreasonable, in retrospect, to assume that such a pre-strained condition became increasingly prevalent in the 25 yr leading up to 1906. Possible seasonal triggering mechanisms may be classified on the basis of how they might affect the three terms in the stress ratio:

$$\frac{\tau}{s_n - p} = K$$

where τ is shear stress, s_n is normal stress, p is pore-fluid pressure along a pre-existing fault, with K the critical ratio above which shear failure will occur in the given medium¹². In theory, on a near-critical fault, any seasonal process which might relatively increase shear stress and/or pore pressure, or relatively decrease normal stress, could potentially destabilize the fault and trigger rupture, and thus account for earthquake seasonality before a great earthquake. Although limited space prevents a detailed discussion of candidate seasonal processes, Table 4 outlines several that might be most effective in the spring. Those listed are ordinarily nominal ambient processes which clearly would not be expected to trigger slip on faults at low strain levels. On faults at high strain levels, however, such as the subparallel faults in the San Andreas system late in a seismic cycle, the listed triggering mechanisms conceivably could introduce a seasonal bias into the time-distribution of earthquakes. In the absence of high pre-stress, such as early in a seismic cycle, or of a seasonal triggering influence, earthquakes would be expected

Table 2 Schuster test results for selected subsets of earthquakes in coastal central California

	1855–1982	1855–1905	1907–1982	1808–1905*	1808–1982*
	≥ 6.0 mag (Independent events)				
<i>n</i>	10	7	2	10	13
ϕ	15 May	30 April	27 August	23–28 May	28 May–2 June
<i>R</i>	3.06	2.48	1.13	5.11–4.93	5.87–5.75
<i>P</i>	0.39	0.42	0.53	0.07–0.09	0.07–0.08
	≥ 5.5 mag (independent events)				
<i>n</i>	27	19	7		
ϕ	30 April	15 April	15 Sept		
<i>R</i>	6.50	8.20	3.37		
<i>P</i>	0.21	0.03	0.20		
	≥ 5.0 mag (independent events)				
<i>n</i>	61	34	26		
ϕ	11 April	15 April	24 October		
<i>R</i>	7.33	9.33	3.05		
<i>P</i>	0.41	0.08	0.70		

n, Number of earthquakes ('annual phases') in subset.

ϕ , Time of year (corresponding to net phase angle) near which earthquakes in subset tended to occur.

R, Vector sum over subset of earthquake phase angles.

P, Probability of observing a vector sum greater than *R* if *n* phases were random (valid for $n \geq 10$)⁵.

* Subset includes three largest pre-1855 central California earthquakes: in 1808, 1836 and 1838 (Table 3). For the undated June 1838 event, extreme dates of 1 June and 30 June were assumed in calculating possible ranges for ϕ and *P*.

ted to distribute themselves randomly throughout the year, as they have been distributed in coastal central California since 1906 (≥ 5.0 mag; Schuster *P* = 0.70).

The above observations provide evidence for precursory seasonal seismicity leading up to the great San Francisco earthquake. Although the observations may be statistically significant, whether they are also physically significant remains to

Table 3 Other strong (≥ 6.0 mag) earthquakes in coastal central and southern California (see text)

Date	Latitude	Longitude	Maximum modified Mercalli intensity	Magnitude
1808 6 21	37°48' N	122°30' W	VIII	?
1836 6 10	37°48' N	122°12' W	VIII	6.8
1838 6 —	37°36' N	122°12' W	VIII	≥ 7.0
1885 4 12	36°24' N	121°00' W	VII	6.2
1890 2 9	33°24' N	116°18' W	VI+	6.3
1892 5 28	33°12' N	116°12' W	V-VI	6.3
1899 7 22	34°18' N	117°30' W	VIII	6.5
1907 9 20	34°12' N	117°06' W		6.0
1910 5 15	33°42' N	117°24' W		6.0
1916 10 23	34°54' N	118°54' W		6.0
1922 3 10	35°48' N	120°18' W		6.5
1927 11 4	34°54' N	120°42' W		7.5
1934 6 8	35°48' N	120°18' W		6.0
1952 11 22	35°42' N	121°12' W		6.0
1983 5 2	36°14' N	120°17' W		6.7

Pre-1900 earthquake data from ref. 7; post-1900 data from ref. 16; 1983 data (Coalinga earthquake) from University of California Seismographic Station, Berkeley. Pre-1900 magnitudes are M_L ; post-1900 magnitudes are M_L , except for 1927 event, which is M_S .

Table 4 Conjectural mechanisms for springtime seismicity

- Pore-fluid pressure increase**
 - Downward-propagating pulse of above average crustal pore pressure due to increased infiltration of water from surface and subsurface reservoirs refilled by winter rainfall.
 - Static unloading of crust in coastal region by seasonal decline in monthly mean sea level of the order of 20 cm, leaving overpressures on pre-stressed faults at depth; minimum occurs in springtime off central California and follows a decline which may exceed 30 cm over a 2-month period (if after a winter-long high sea level anomaly, or 'El Nino')¹⁷⁻¹⁹.
- Increase in shear stress and/or decrease in normal stress**
 - Springtime increase in northwestward force vector on northwestward-moving Pacific plate due to increasing inclination of Northern Hemisphere toward Sun. (As this notion might predict, strong earthquakes on the 1,100-km long SAF system tend to occur near a date earlier in the spring in the south (31 March) than in the north (late May).)
 - Interaction at the plate boundary of torques on the Pacific and North American plates resulting from the springtime deceleration of Earth's rotational spin (manifested by springtime increase in the length of day)²⁰.

be demonstrated. Nevertheless, the evidence corroborates, if rather oddly, the suspicion that a seismic cycle governs the long-term behaviour of the northern SAF system. If the historical pattern of earthquake seasonality is characteristic of that long-term behaviour, several interesting inferences are possible.

(1) Based on the absence since 1906 of frequent spring earthquakes of ≥ 5.5 mag in coastal central California, (or, perhaps, of high seismicity in another season), a repeat of the 1906 earthquake on the San Andreas fault is probably at least a quarter-century away (as others have concluded¹). (2) Should a seasonal seismicity pattern re-emerge, the contemporary practice of estimating cumulative and annual probabilities of expected major earthquakes (see, for example, refs 11, 13) may be refined to the precision of seasonal probabilities. (3) Seasonal analysis of historical seismicity on shallow faults elsewhere in the world may yield similarly predictive results. (4) Seasonal analysis of other kinds of solid-Earth geophysical data, such as ground tilt, strain, creep, gravity, tectonomagnetism, microseismicity, well-water level and frequent data from land-based and astrometric geodesy obtained by radio and laser interferometric surveys, may likewise identify informative relationships to earthquake activity. In this manner, some seasonal variations

in instrumental data hitherto attributed to meteorological 'noise' may prove to have a deeper significance.

I thank W. Thatcher, A. G. Lindh, R. Wesson, W. H. Prescott and D. H. Peterson for discussions, T. R. Bruns for the loan of his calculator, and F. W. Klein and J. S. Derr for assistance in the preparation of the manuscript. I also thank D. S. McCulloch, G. W. Hill, J. S. Derr and W. L. Ellsworth for their support. Ideas and opinions expressed herein are the sole responsibility of the author, except where otherwise cited, and do not represent official views and conclusions of the US Geological Survey or the Minerals Management Service.

Received 13 May; accepted 31 October 1983.

- Ellsworth, W. L., Lindh, A. G., Prescott, W. H. & Herd, D. G. in *Earthquake prediction—An International Review* (eds Simpson, D. W. & Richards, P. G.) 126–140 (American Geophysical Union, Washington DC, 1981).
- Schuster, A. *Proc. R. Soc.* **61**, 455–564 (1897).
- Shlien, S. *Geophys. J. R. astr. Soc.* **28**, 27–34 (1972).
- Heaton, T. H. *Geophys. J. R. astr. Soc.* **43**, 307–326 (1975).
- Klein, F. W. *Geophys. J. R. astr. Soc.* **45**, 245–295 (1976).
- Kilston, S. & Knopoff, L. *Nature* **304**, 21–25 (1983).
- Topozada, T. R., Real, C. R. & Parke, D. L. *Calif. Div. Mines Geol. Open File Rep.* 81–11 SAC (1981).
- Hall, N. T., Cotton, W. R. & Hay, E. A. in *AGU Chapman Conf. on Fault Behaviour and the Earthquake Generation Process* Abstr. (American Geophysical Union, Washington DC, 1982).
- Fedotov, S. A. *Akad. Nauk. SSSR Inst. Fiz. Zeml.: Trudy* **36**, 66–93 (1968).
- Mogi, K. in *Earthquake Prediction—An International Review* (eds Simpson, D. W. & Richards, P. G.) 43–51 (American Geophysical Union, Washington DC, 1981).
- Raleigh, C. B., Sieh, K. E., Sykes, L. R. & Anderson, D. L. *Science* **217**, 1097–1104 (1982).
- Raleigh, C. B. in *Underground Waste Management and Environmental Implications* Mem. 18 (ed. Cook, T. D.) 273–279 (American Association of Petroleum Geologists, Tulsa, Oklahoma, 1972).
- U.S. geol. Surv. Open-File Rep.* 81–115 (1981).
- Topozada, T. R. *Bull. seism. Soc. Am.* **65**, 1223–1238 (1975).
- Jennings, P. & Kanamori, H. *Earthquake Inform. Bull.* **13**, 49–51 (1981).
- Topozada, T. R., Real, C. R. & Parke, D. L. *California Division of Mines and Geology Map Sheet* 39 (1978).
- Pattullo, J., Munk, W., Revelle, R. & Strong, E. *J. mar. Res.* **14**, 88–155 Appendix I and Chart 1 (1955).
- Enfield, D. B. & Allen, J. S. *J. phys. Oceanogr.* **10**, 557–578 (1980).
- Price, J. M. *J. phys. Oceanogr.* **11**, 1375–1302 (1981).
- Lambeck, K. & Hopgood, P. *Geophys. J. R. astr. Soc.* **71**, 581–587 (1982).

Localized uplift in the Scottish Dalradian

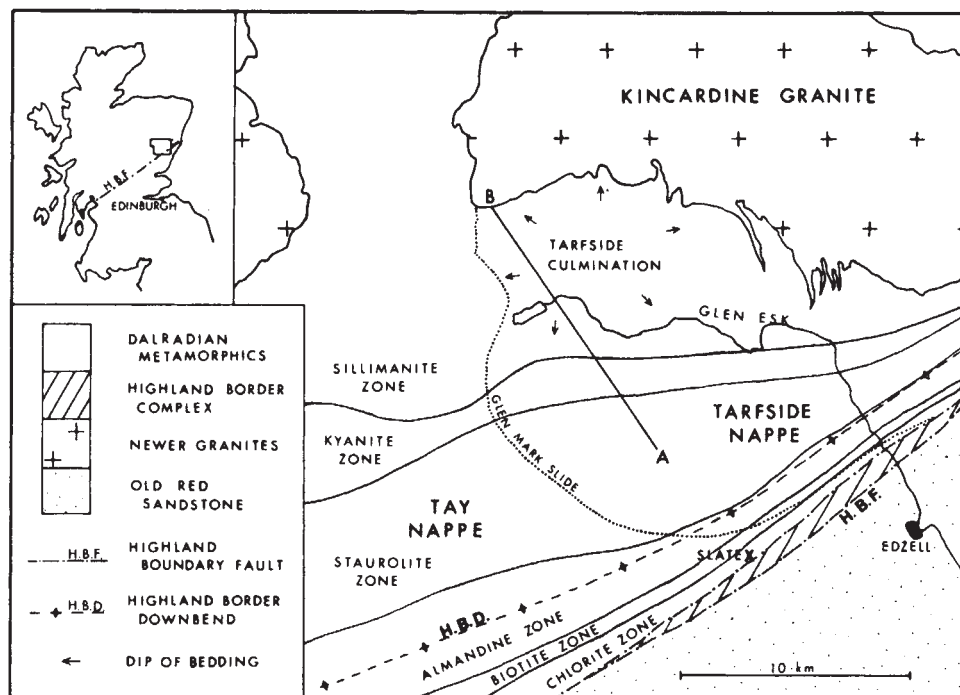
T. J. Dempster

Grant Institute of Geology, University of Edinburgh,
West Mains Road, Edinburgh EH9 3JW, UK

Rb-Sr mica ages are reported here, from the Dalradian meta-sediments of the Angus region in the Scottish Highlands (Fig. 1), in a study of the uplift history of this classic metamorphic terrain. The samples come from across the Barrovian metamorphic zones¹, and provide evidence of diachroneity in both metamorphism and later cooling. Thus, the oldest mineral ages, and hence earliest cooling, are found in the highest grade metamorphic rocks to the north. The results show the importance of tectonically-controlled, localized uplift processes. Given that erosion is an important influence on the thermal development of metamorphic rocks², the constant or steadily decaying isostatic uplift rates, often assumed in models of orogenic belts^{2,3}, may be a marked oversimplification of the real events.

The major structure of the Dalradian is dominated by pre-metamorphic nappe structures⁴, represented by the Tay and Tarfside Nappes in the present area⁵ (Fig. 1). The rocks considered in this study come from the roughly flat-lying, right way-up limb of the Tarfside Nappe, along the line A–B in Fig. 1. Essentially, the area is considered to show a continuation of the flat belt and Highland Border Steep Belt of Perthshire, with the axial trace of the post-metamorphic Highland Border Down-bend dividing these two domains. The flat belt region, in the area of the A–B traverse, also shows the complication of the late Tarfside Culmination⁶ in the north of the area (Fig. 1). Regional metamorphic grade increases away from the Highland Boundary Fault, from the chlorite zone⁷ to the sillimanite zone in the north¹. Estimates of the age of peak metamorphism in the Dalradian, range from 515 to 490 Myr (refs 8, 9), although

Fig. 1 Structural and metamorphic map of the Angus region in the Scottish Highlands (partly based on Barrow¹ and Harte²), showing the location of traverse A-B.



the Lower Palaeozoic metamorphism over the Scottish Highlands as a whole may show considerable diachronism^{10,11}. Metamorphism in this traverse is roughly contemporaneous with other areas in the Dalradian, with the metamorphic climax approximately synchronous with the third regional deformation event (D3). However, slightly earlier garnet and staurolite porphyroblast growth, relative to the D3 mica fabric, is seen in the lower grade rocks, in the south of the traverse.

The interpretation of isotopic mineral ages in slowly cooling metamorphosed rocks is largely based on the theory of closure temperatures^{12,13}. This predicts that minerals will record the times at which they cool through characteristic temperatures, below which the mineral is closed to the relevant isotopic exchange. Thus, biotites, with approximate Rb-Sr closure temperatures of 300 °C (ref. 12), record younger ages than coexisting muscovites, with higher Rb-Sr closure temperatures of about 500 °C (ref. 12). Dewey and Pankhurst¹⁴ examined K-Ar mica ages from the Scottish Highlands and showed a broad relationship between the mineral ages and metamorphic grade, with younger ages and delayed cooling in the hotter, more deeply buried rocks. Thus, it might be expected that with roughly simultaneous metamorphism and uniform uplift, earlier cooling to mica closure temperatures, and hence older ages, would be seen in the lower grades to the south, along the present profile.

In this detailed study along the traverse A-B (Fig. 1), biotite and muscovite were separated from biotite-muscovite-plagioclase-bearing schists and analysed, together with whole-rock samples by Rb-Sr isotope dilution methods described elsewhere¹⁵. Details of all mineral age results, presented in Fig. 2, are available on request, and a selection of isotopic data is given in Table 1. The grain sizes and chemical compositions of each mica are relatively constant in the samples of this study and hence probably have little control on closure temperatures¹⁷ or age variation. Thus, the isotopic age pattern (Fig. 2) largely reflects variation in the time of cooling to approximately uniform closure temperatures. It should be emphasized that the relative mica ages observed are more critical to any interpretation than the absolute ages and values of closure temperatures adopted.

The Rb-Sr mineral ages are corrected for initial Sr using the whole-rock analysis, and if the mica ages in the narrow thermal aureole of the Kincardine Granite are examined, this approach seems to be justified. Thus, mica ages are reset to an age of intrusion (~410 Myr), consistent with other independent estimates of the timing of Newer Granite emplacement¹⁸. So, the micas must exchange with a low ⁸⁷Sr/⁸⁶Sr phase, and any uncer-

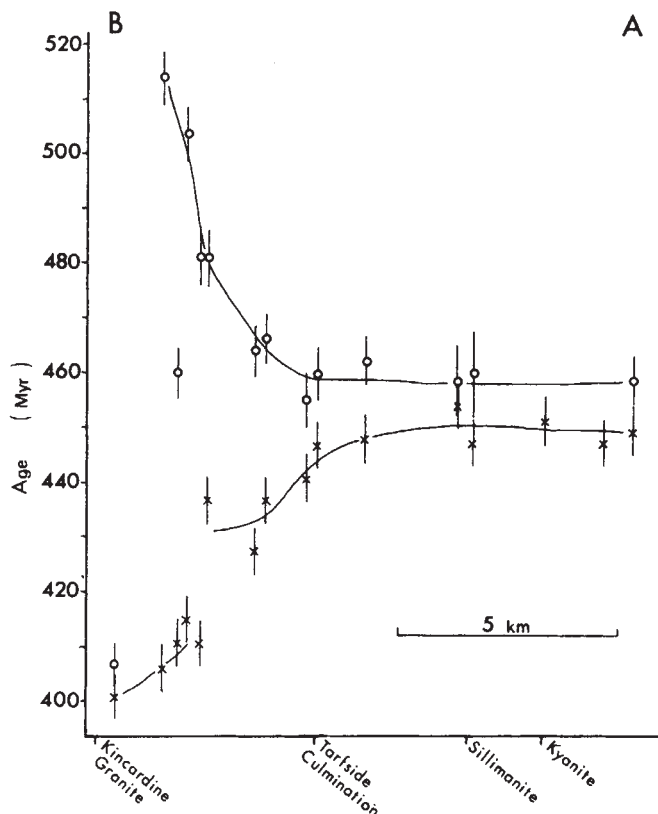


Fig. 2 Mineral age results from traverse A-B shown in Fig. 1. ○, Rb/Sr muscovite; ×, Rb/Sr biotite.

tainty in the exact composition of this phase, is more than outweighed by the large spread in Rb-Sr ratios of the exchanging phases.

The whole-rock K-Ar dating of a well-crystallized slate, containing muscovite as its only major K-bearing phase and only one generation of white-mica, gives an estimate of the timing of metamorphism⁹ in the chlorite zone (Fig. 1) of 489 ± 10 Myr (Table 1). Surprisingly, the oldest Rb-Sr muscovite age of 514 ± 5 Myr is found in the north of the traverse, which gives a minimum estimate of the timing of metamorphism in the high grades and supports textural evidence of diachroneity of

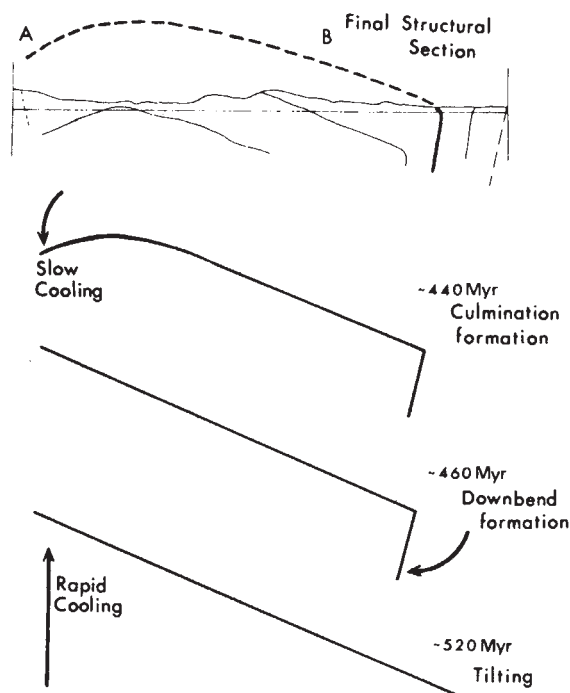


Fig. 3 Schematic sections showing the proposed structural evolution and cooling history that results in the mineral ages shown in Fig. 2.

metamorphism in the region. The old muscovite cooling ages in the north of the transect (Fig. 2) must indicate extremely rapid cooling, from the estimated peak metamorphic temperatures of $\sim 650^\circ\text{C}$ (ref. 18), to below the approximate Rb–Sr closure temperatures of muscovite of $\sim 500^\circ\text{C}$ (ref. 12) (although this rapid cooling itself may slightly increase the closure temperatures¹⁶). No evidence of this early rapid cooling is seen immediately to the south, where, despite a roughly similar (D3) textural age of metamorphism, initial slow cooling is indicated by Rb–Sr muscovite ages of ~ 460 Myr. However, a gradual transition occurs from the north towards the centre and to the south of the Tarfside Culmination occurs, as shown by the steady southerly decrease in the muscovite ages. After the initial rapid cooling in the northern area, very slow cooling must have followed as shown by the relatively young biotite Rb–Sr ages

of ~ 430 Myr. Again contrasting behaviour is seen to the south of the Tarfside Culmination, in that the similarity of muscovite (~ 460 Myr) and biotite (~ 450 Myr) ages implies relatively rapid cooling and uplift for this area at a late stage.

Neither of the prominent post-metamorphic structures, the Highland Border Downbend or the Tarfside Culmination, seem to be capable of causing the early uplift, which is necessary for the rapid cooling in the very north. The Downbend structure must postdate the cooling and uplift in the sillimanite zone, as it deforms the Arenig fossil-bearing¹⁹ (~ 480 Myr)²⁰ Highland Border Complex²¹, and also folds the metamorphic zones of the lower grade areas¹⁸, which are developed after cooling in the north. Similarly, the Tarfside Culmination is a young dome-structure and the old ages come from the northwestern rather than the central parts of the dome (Fig. 2). A striking feature of the Angus region is the significant dip of the 'flat belt' (to the south of the Tarfside Culmination) at angles of $\sim 30^\circ$ to the south-east, compared with the virtually horizontal beds of Perthshire⁴. It is suggested, as shown in Fig. 3, that an early tilting of originally flat-lying beds could result in rapid uplift, erosion and cooling in the most northerly rocks, which before the Culmination formation would be the highest level rocks (Fig. 3). This suggests that the Culmination forms subsequently by relative downwarping of the beds to the north, as a consequence of which, the once coolest rocks to the north, become the slowest cooling and show the youngest biotite ages (Fig. 2).

Thus, tectonic activity may be largely responsible for the cooling and uplift pattern in this part of the Highlands. A syn- to post-metamorphic tilting would also explain the diachronous metamorphism in this area, with convection of heat by rapid uplift², hastening the approach to peak metamorphic conditions. Thus, earlier-metamorphosed, higher-temperature rocks will be exposed in more rapidly uplifted areas. The approximately horizontal dip of the flat belt elsewhere in the Southern Dalradian, suggests that not only is the rapid early uplift localized in the north of the area considered, but is probably not seen elsewhere along the belt. This may account for the exposure of low level Tarfside Nappe rocks at the expense of the inverted Tay Nappe rocks to the south-west⁵ (Fig. 1).

Clearly, such major cooling and uplift rate variations cannot be accommodated into regional (on the scale of the Grampian Highlands) block uplift events with steady isostatic return to normal crustal thicknesses^{2,22}. It is apparent that localized uplift relating to structural activity and possibly decoupling and movement of small blocks in the basement is an important late tectonic

Table 1 Analytical data

Sample	Rb (p.p.m.)	Sr (p.p.m.)	$^{87}\text{Rb}/^{86}\text{Sr}$ (atomic)	$^{87}\text{Sr}/^{86}\text{Sr} \pm 2\sigma$ (atomic)	Rb–Sr mineral age (Myr)
GL 771					
Whole-rock	97.8	158.6	1.7882	0.73450 ± 20	
Muscovite	183.2	86.2	6.1847	0.76332 ± 30	460 ± 8
Biotite	342.4	30.9	32.721	0.93166 ± 50	447 ± 4
GL 813					
Whole-rock	63.1	126.6	1.4460	0.74034 ± 60	
Muscovite	240.9	31.7	22.388	0.87942 ± 30	466 ± 5
Biotite	484.3	9.4	163.89	1.7506 ± 4	437 ± 4
GL 831					
Whole-rock	114.7	75.4	4.4275	0.75899 ± 20	
Muscovite	250.3	47.9	15.299	0.83865 ± 30	514 ± 5
Biotite	475.5	9.0	167.43	1.7024 ± 2	406 ± 4
GL 841					
Whole-rock	84.4	46.0	5.3377	0.76528 ± 4	
Muscovite	260.0	16.9	45.849	0.99995 ± 8	406 ± 4
Biotite	510.4	7.1	236.12	2.0838 ± 4	401 ± 4
	K (wt%)	$^{40}\text{Ar}_{\text{rad}}$ (p.p.m.)	%Rad	K–Ar Age (Myr)	
GL 691					
Whole-rock Slate	3.46	0.13473	98.2	489 ± 10	

Decay constants and isotopic abundances of Steiger and Jäger²⁵ are used.

feature of the Dalradian evolution. Given that there seems little reason to believe that the Lower Palaeozoic metamorphic belt of the Scottish Highlands should be markedly different from other orogenic belts, thermal models of metamorphic evolution in thickened crust, involving steadily decaying uplift processes^{3,22}, are probably an oversimplification.

These results must also cast doubt on the technique of assessing closure temperature variation, by comparing mineral ages over large tracts of metamorphic rocks on the assumption of uniform uplift and metamorphism²³. The occurrence of localized uplift phenomena, within a regional metamorphic terrain, will also influence the variation in metamorphic grade or facies series or piezothermic arrays²⁴ (metamorphic geotherms²) across the terrain. Thus, a change in facies series¹⁸ may partly reflect localized lateral (rather than vertical) differences in *P-T*-time paths for individual rock units. One further consequence of localized uplift is that it may also cause lateral redistribution of heat, thus, small uplifting areas may act like syn-metamorphic intrusions with respect to surrounding non-uplifting areas.

The support of a NERC grant is acknowledged and I thank Ben Harte for his encouragement and supervision, and John Hutchinson for assistance with mass spectrometry performed at the Scottish Universities Research and Reactor Centre.

Received 12 July; accepted 4 October 1983.

- Barrow, G. *Proc. geol. Ass.* **23**, 268–290 (1912).
- England, P. C. & Richardson, S. W. *J. geol. Soc. Lond.* **136**, 489–495 (1977).
- England, P. C. *Tectonophysics* **46**, 21–40 (1978).
- Bradbury, H. J., Harris, A. L. & Smith, R. A. in *The Caledonides of the British Isles—Reviewed* (eds Harris, A. L., Holland, C. H. & Leake, B. E.) 213–220 (Scottish Academic, Edinburgh, 1979).
- Harte, B. in *The Caledonides of the British Isles—Reviewed* (eds Harris, A. L., Holland, C. H. & Leake, B. E.) 221–228 (Scottish Academic, Edinburgh, 1979).
- Read, H. H. *Trans. R. Soc. Edinb.* **55**, 755–772 (1928).
- Tilley, C. E. *J. geol. Soc. Lond.* **81**, 100–112 (1925).
- Pankhurst, R. J. & Pidgeon, R. T. *Earth planet. Sci. Lett.* **22**, 256–266 (1974).
- Harper, C. T. *Scott. J. Geol.* **3**, 46–66 (1967).
- van Breemen, O., Aftalion, M., Pankhurst, R. J. & Richardson, S. W. *Scott. J. Geol.* **15**, 49–62 (1979).
- van Breemen, O. & Piasecki, M. A. J. *J. geol. Soc. Lond.* **140**, 47–62 (1983).
- Purdy, J. W. & Jäger, E. *Mem. Instn. Geol. Miner. Univ. Padova* **30**, 1–32 (1976).
- Desmons, J., Hunziker, J. C. & Delaloye, M. *Contr. Miner. Petrol.* **80**, 386–390 (1982).
- Dewey, J. F. & Pankhurst, R. J. *Trans. R. Soc. Edinb.* **68**, 361–389 (1970).
- Blaxland, A. B., van Breemen, O., Emeleus, C. H. & Anderson, J. G. *Bull. geol. Soc. Am.* **89**, 231–244 (1978).
- Dodson, M. H. *Contr. Miner. Petrol.* **40**, 259–274 (1973).
- Pidgeon, R. T. & Aftalion, M. *Geol. J. Spec. Iss.* **10**, 183–220 (1978).
- Harte, B. & Hudson, N. F. C. in *The Caledonides of the British Isles—Reviewed* (eds Harris, A. L., Holland, C. H. & Leake, B. E.) 323–337 (Scottish Academic, Edinburgh, 1979).
- Curry, G. B., Ingham, J. K., Bluck, B. J. & Williams, A. J. *J. geol. Soc. Lond.* **139**, 451–454 (1982).
- Bluck, B. J., Halliday, A. N., Aftalion, M. & Macintyre, R. M. *Geology* **8**, 492–495 (1980).
- Booth, J. E. thesis, Univ. Edinburgh (1983).
- Wells, P. R. A. *J. geol. Soc. Lond.* **136**, 663–671 (1979).
- O'Nions, R. K., Smith, D. G. W., Baadsgaard, H. & Morton, R. D. *Earth planet. Sci. Lett.* **5**, 339–345 (1969).
- Richardson, S. W. & England, P. C. *Earth planet. Sci. Lett.* **42**, 183–190 (1979).
- Steiger, R. H. & Jäger, E. *Earth planet. Sci. Lett.* **36**, 359–362 (1977).

Period multippling-evidence for nonlinear behaviour of the canine heart

Amy L. Ritzenberg*, Dan R. Adam†
& Richard Jonathan Cohen*†

* MIT Department of Physics and † Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Although there has recently been considerable interest in applying the theory of nonlinear dynamics to the analysis of complex systems^{1,2}, as yet applications of the theory to biological systems *in vivo* have been very limited. We report here evidence of nonlinear behaviour in the electrocardiogram and arterial blood pressure traces of the noradrenaline-treated dog. Noradrenaline produces variations in these traces that repeat themselves with regular periods of integral numbers of heart-beats (period multippling), an effect that resembles the 'period-doubling' and other 'bifurcative' behaviour^{3–5} observed when the driving

frequency of a nonlinear oscillator is increased above a critical value^{6–9}. The simplest type of periodic variation that we report is the so-called 'electrical alternans', which has long been known as one response of cardiac electrical activity to certain stresses and disease states^{10–13}.

Ten dogs were anaesthetized with sodium pentobarbitol and ventilated with a respirator at a fixed rate (to distinguish electrocardiogram (ECG) modulation due to respiration). Saline-filled catheters were introduced into the femoral artery and vein, and arterial blood pressure (ABP) tracings obtained. Unipolar leads were placed precordially and the surface ECG recorded. Noradrenaline bitartrate was given intravenously in a dosage sufficient to intoxicate (indicators being substantial sinus tachycardia and arterial hypertension). Dosages ranged from 0.02 to 0.05 mg kg⁻¹. To confirm the presence of period multippling more rigorously than via a simple visual inspection of the traces, ECG and ABP records were digitized and then fast Fourier transformed and the spectra plotted. According to Fourier theory, the Fourier spectrum of a record with no period multippling should contain sharp peaks only at the heart rate and at integral multiples (the Fourier harmonics) of that frequency. Period doubling (alternans) should appear in the spectrum as an isolated peak (subharmonic) at a frequency equal to one-half the heart rate and additional peaks at integral multiples of this frequency; period tripling should appear as peaks at integral multiples of the frequency equal to one-third the heart rate, and so on (hence 'period multippling' corresponds to frequency dividing).

The results of this means of analysis are summarized in Table 1. One of the 10 dogs fibrillated spontaneously upon injection with noradrenaline and is thus excluded from the study. Of the nine remaining, seven displayed modulations of clearly discernible order. There were several instances of period doubling, tripling, quadrupling and quintupling. Figure 1 shows typical episodes of tripling and quadrupling; ECG, ABP and their respective Fourier spectra are displayed. The time scale for these observations is as follows: episodes of 4.25 s or more were considered stable. Intervals of apparent stability ranged from approximately 6 s (an instance of quadrupling) to tens of seconds (instances of doubling). Inter-episode (apparently aperiodic) times varied from tens of seconds to minutes. In this study, multippling was always found in both ECG and ABP, but never in only one. Thus, we would expect that this phenomenon is not only electrical (periodic alterations of the pattern of conduction through the heart) but mechanical as well (affecting the contractile action of the heart as a whole). Since the frequency of the ECG modulation was unrelated to the frequency of respiration, the effect is not respiratory in origin.

If the analogy with a driven, nonlinear oscillator is appropriate, the heart rate should correlate strongly with the period multippling regime¹⁴. For example, in our experience of oscillators doubling precedes either tripling or quadrupling, quadrupling precedes quintupling, but quadrupling precedes tripling, due to the particular way in which the new regimes are generated (beyond the scope of this letter), as the driving frequency of

Table 1 Period multippling regimes observed

Animal	Regime	Ventricular rate	Animal	Regime	Ventricular rate
1	Doub.	211	7	Quad.	183
	Quad.	228		Trip.	198
2	Doub.	253	9	Quin.	225
	Quad.	267		Doub.	105
3	Doub.	225	10	Trip.	112
	Quad.	253		Doub.	112
4	Doub.	190			
	Trip.	295			

Experimental animals are numbered 1–10. For the seven animals in whose ECG and ABP traces regimes of period multippling were observed, the various regimes are displayed.

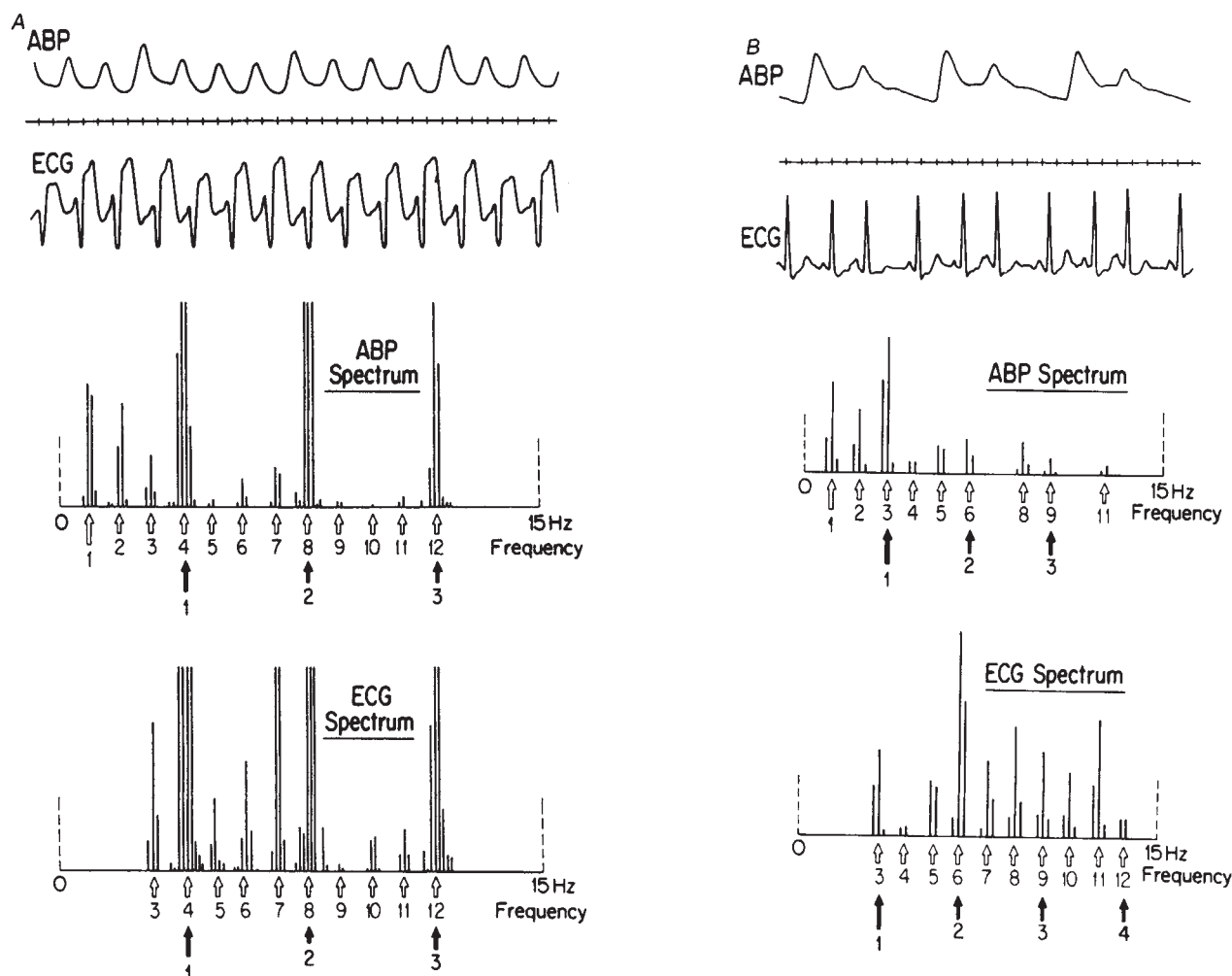


Fig. 1 ABP and ECG traces are displayed, followed by their respective fast Fourier transforms. In *A*, the heart rate is 228 beats per min, making the beat frequency 3.8 Hz (large black arrow). Subharmonics (large white arrow for the first, at 0.95 Hz) appear at quarters of the beat frequency, indicating quadrupling. In *B*, the heart rate is 198 beats per min making the fundamental frequency of atrial beating 3.3 Hz. This appears (large black arrow) as a major peak in both spectra. Subharmonics (large white arrow for the first, at 1.1 Hz) appear at thirds of this frequency, indicating period tripling. In both *A* and *B*, higher Fourier harmonics are shown with smaller arrows; black for higher harmonics of the beat frequency and white for higher harmonics of the multiplying frequency. The number below the arrow gives the order of harmonic. (The amplitudes of the first few subharmonics in the ECG spectra are too low to appear on these plots; their higher Fourier harmonics show their presence.)

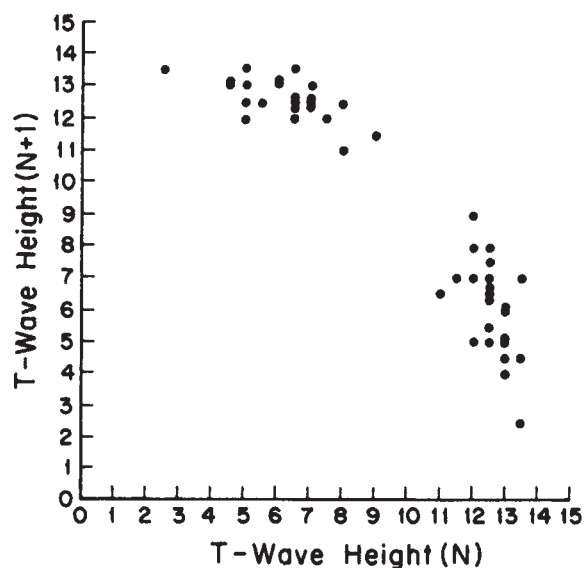


Fig. 2 The height of the $N+1$ st T-wave plotted against the height of the N th T-wave during 48 beats of an episode of period doubling (animal 1).

the nonlinear oscillator is raised. This precedence of period multiplying regimes with increasing heart rate was obeyed in all cases (Table 1).

The period multiplying behaviour described here may arise in different instances from different electrophysiological mechanisms. For example, in Fig. 1A, the intraventricular beat interval is fixed. Period quadrupling in the ECG record is comprised of modulation of the shape of the QRS and T-wave complexes. This suggests that the electrical pattern of depolarization and repolarization varies on a beat-to-beat basis. In Fig. 1B, the period tripling shown in ECG and ABP spectra reflects the timing of atrial excitations (which effectively penetrate the atrioventricular junction to activate the ventricles) as well as the sort of excitation pattern variability of Fig. 1A. Other mechanisms of period multiplying include medically well known bigeminy, trigeminy and quadrigeminy, in which every second, third or fourth ventricular beat originates from a ventricular focus; and a variety of forms of second-degree atrioventricular block. We suggest here that period multiplying is a common response of the electrical and mechanical activity of the stressed heart, even when apparently diverse mechanisms are operative.

The observation of period multiplying in the canine heart upholds the naive expectation that this system is nonlinear, since a strictly linear system can never exhibit period doubling^{3,14}. Recent work on the driving and resetting of neural oscillators,

including excised heart tissue, supports this¹⁵⁻¹⁹. One simple way to analyse such a system for linearity is to explore the functional dependence of the shape of one beat on the previous beat. We have generated 'scatter diagrams' (Fig. 2) of the T-wave peak height or T-wave area for the $N+1$ st beat against the N th. The degree of 'scatter' (range of ordinates corresponding to one abscissa) is a measure of how closely the $N+1$ st beat depends for its shape solely on the shape of the N th. Our plots show both the strong causal link between the N th T-wave amplitude and the $N+1$ st (that is, a small degree of scatter) and the nonlinearity of that causal relationship.

Further, from Li and Yorke's proof²⁰ the existence of tripling indicates that the cardiac system is sufficiently nonlinear to produce responses to atrial driving of all periodicities plus pseudostochastic sequences of waveforms. This leads us to ask why we saw no orders of period multippling higher than that of quintupling. A related question is why signals of a relatively low order of multippling alternate chronologically in our records (corresponding to similar heart rates and other system parameters) with aperiodic stretches. One answer is that there are several scenarios for the period multippling of noiseless nonlinear oscillators which predict the existence of intercalated bands of chaotic and periodic activity; in particular, the phenomenon of

'intermittency' is relevant^{21,22}. Alternatively, noise disrupts the bifurcation structure of an idealized, non-linear system^{8,9}. Higher order orbits are empirically more fragile (due to their smaller parameter space) and subject to effacement, so they are simply less likely to be observed. For example, Duffing's oscillator, when subjected to white noise, is found to develop a symmetrical gap in its bifurcation sequence. Higher-order orbits can even be induced solely by increasing the amplitude of the noise. In the heart, there are many sources of noise, in the form of stochastically varying parameters governing cardiac function; such as intrinsic fluctuations in the atrial rate, and variation in conduction and contractility due to local areas of ischaemia that develop in the stressed heart.

Our results support the model of the noradrenaline intoxicated heart as a highly nonlinear, driven oscillator subject to noise thus usefully linking many seemingly disparate dysrhythmias as bifurcative responses to variation in system parameters.

This research was supported by NSF grants 7922091-ECS and 8121571-ECS, ONR grant N00014-80-C-0520, a grant from the Whitaker Health Sciences Fund, and a grant from the R. J. Reynolds Industries. A.L.R. received support from a Whitaker Health Sciences Fund Fellowship and R.J.C. from a Hartford Foundation Fellowship.

Received 26 July; accepted 14 September 1983.

1. Wolf, A. *Nature* **305**, 182-183 (1983).
2. Holden, A. V. *Nature* **305**, 183 (1983).
3. May, R. & Oster, G. *Am. Nat.* **110**, 573-599 (1976).
4. Feigenbaum, M. J. *J. statist. Phys.* **19**, 25-52 (1978).
5. Hofstadter, D. R. *Sci. Am.* **245**, 22-43 (1981).
6. Huberman, B. A. & Crutchfield, J. P. *Phys. Rev. Lett.* **43**, 1743-1747 (1979).
7. Guevara, M. R. & Glass, L. *J. math. Biol.* **14**, 1-23 (1982).
8. Glass, L., Graves, C., Pettillo, G. & Mackey, M. J. *theor. Biol.* **86**, 455-475 (1980).
9. Huberman, B. A. & Crutchfield, J. P. *Phys. Lett. A* **77**, 407-410 (1980).
10. Hellerstein, H. K. & Liebow, I. M. *Am. Heart J.* **160**, 366-374 (1949).

11. Hoffman, B. F. & Suckling, E. E. *Am. J. Physiol.* **179**, 123-130 (1954).
12. Kleinfield, M. & Stein, E. *Am. Heart J.* **75**, 528-530 (1968).
13. Adam, D. R., Akselrod, S. & Cohen, R. J. *Comp. Cardiol.* **8**, 307-310 (1981).
14. Feigenbaum, M. J. *Los Alamos Sci.*, 2-27 (1980).
15. Mouloupoulos, S. D., Kardaras, N. & Sideris, D. A. *Am. J. Physiol.* **208**, 154-157 (1965).
16. Winfree, A. T. *Science* **197**, 761-762 (1977).
17. Moe, G., Jalife, J., Mueller, W. & Moe, B. *Circulation* **56**, 968-979 (1977).
18. Jalife, J. & Antzelevitch, C. *Science* **206**, 695-697 (1979).
19. Guevara, M., Glass, L. & Shrier, A. *Science* **214**, 1350-1353 (1981).
20. Li, T. & Yorke, J. *Am. math. Mon.* **82**, 985-992 (1975).
21. Metropolis, N., Stein, M. L. & Stein, P. R. *J. Comb. Theory* **A15**, 25-43 (1973).
22. Eckmann, J. P. *Rev. mod. Phys.* **53**, No. 4, Pt 1, 643-654 (1981).

Brain potentials during reading reflect word expectancy and semantic association

Marta Kutas & Steven A. Hillyard

Department of Neurosciences, M-008, University of California, La Jolla, California 92093, USA

The neuroelectric activity of the human brain that accompanies linguistic processing can be studied through recordings of event-related potentials (e.r.p. components) from the scalp. The e.r.ps triggered by verbal stimuli have been related to several different aspects of language processing¹. For example, the N400 component, peaking around 400 ms post-stimulus, appears to be a sensitive indicator of the semantic relationship between a word and the context in which it occurs. Words that complete sentences in a nonsensical fashion elicit much larger N400 waves than do semantically appropriate words or non-semantic irregularities in a text^{2,3}. In the present study, e.r.ps were recorded in response to words that completed meaningful sentences. The amplitude of the N400 component of the e.r.p. was found to be an inverse function of the subject's expectancy for the terminal word as measured by its 'Cloze probability'. In addition, unexpected words that were semantically related to highly expected words elicited lower N400 amplitudes. These findings suggest N400 may reflect processes of semantic priming or activation.

Late negative e.r.ps resembling the N400 have been observed in experiments that required subjects to make decisions about words based on their semantic attributes⁴⁻⁷ and Stuss *et al.*⁸ noted an N400-like component following isolated, single words or pictures that required naming. One likely interpretation of these findings would assume that N400 amplitude reflects the extent to which a word is unpredictable or unexpected, regardless of whether or not it is incongruous with a preceding context.

Since word expectancy influences how rapidly and accurately words are accessed, recognized and understood⁹, a physiological index of this process would have considerable usefulness for revealing the structure of language comprehension mechanisms.

We examined this relationship by recording e.r.ps to words that completed sentences in a meaningful way but varied systematically in the degree to which they were expected. The sentences were selected from a set in which the degree of expectancy for alternative terminal words had been determined using the 'Cloze' procedure; that is, by requiring a large group of subjects to fill in the missing terminal word¹⁰. A word's Cloze probability is defined as the proportion of subjects using that word to complete a particular sentence. The experimental design called for words having different Cloze probabilities (hi, med, or lo) to be placed at the ends of sentences having one of three levels of contextual constraint (hi, med, or lo). Highly constrained sentences were those that led to very predictable endings, while sentences of low constraint did not induce such strong expectations (for examples see Fig. 1A).

A total of 321 sentences were presented, one word at a time, on a video terminal controlled by a microcomputer. Words were presented once every 700 ms for a duration of 132 ms. Fourteen subjects were instructed to read the sentences silently in order to answer a questionnaire about their contents at the end of the experiment. Scalp electrical activity was recorded using non-polarizable electrodes from frontal (Fz), central (Cz), parietal (Pz), and occipital (Oz) midline locations and from symmetrical sites over the anterior temporal (AT) and posterior temporal (PT) regions of the left (L) and right (R) hemispheres, each referred to linked mastoids. Eye movements and blinks were monitored via infraorbital and external canthal electrodes. The midline recordings were amplified with a bandpass of 0-40 Hz; the lateral recordings and electrooculogram with a 0.01-40 Hz bandpass.

The e.r.p. waveforms in Fig. 1B show that highly probable words at the ends of highly constrained sentences were followed by a broad, late positivity (hi/hi, solid tracing). In contrast, the low probability words elicited a posteriorly distributed negative

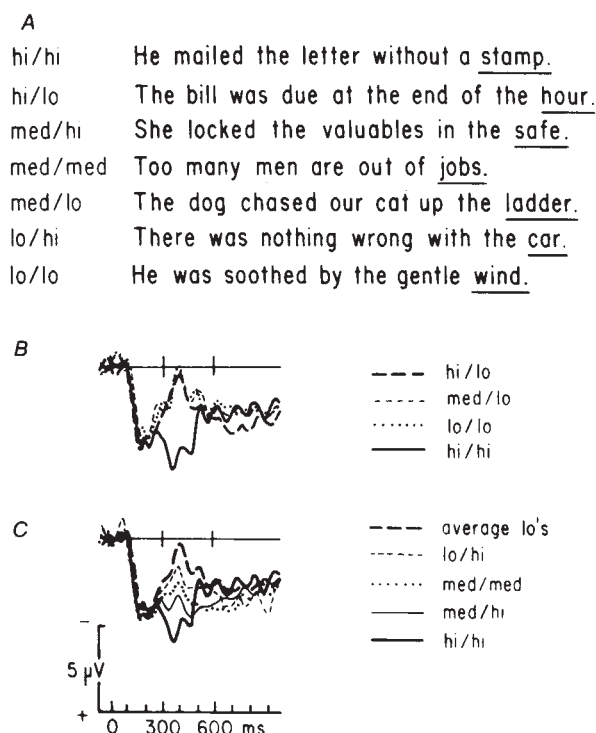


Fig. 1 **A**, An example from each of seven classes of sentences that varied in degree of contextual constraint and in the Cloze probability of the terminal word. There were 40–50 of each sentence type. On the left are shown the levels of contextual constraint/Cloze probability for each class. These two dimensions are not wholly independent, since only the more highly constrained sentences have the possibility of being terminated by words of very high Cloze probability. Thus, the level of Cloze probability considered to be 'hi' was relative to the level of contextual constraint, having average values of 0.92, 0.63 and 0.29 for the hi/hi, med/hi and lo/hi sentences, respectively. All of the low Cloze probability sentences, however, had average values of less than 0.03. **B**, Superimposition of grand average e.r.ps (across 14 subjects) from the Pz electrode to the low Cloze probability words terminating sentences of high, medium and low contextual constraint, together with the e.r.p. to high Cloze probability words completing highly constrained sentences. **C**, Grand average e.r.ps to terminal words of varying Cloze probability terminating sentences of high, medium and low contextual constraint. The e.r.p. to low Cloze probability words was averaged across the three levels of contextual constraint shown in **B**.

component (N400) that was superimposed upon the positive shift. The N400 amplitude to the low probability endings did not vary significantly over the three levels of contextual constraint (compare dotted and dashed tracings).

The N400 was measured as the mean amplitude over 300–500 ms post-stimulus, relative to a 50 ms prestimulus baseline. In general, the N400 amplitude was more sensitive to Cloze probability than to the degree of contextual constraint. For example, the e.r.ps to lo/hi versus med/med words, which were very similar in average Cloze probability (0.29 versus 0.23) but completed sentence fragments of low and medium constraint, respectively, did not differ significantly in N400 amplitude. On the other hand, comparisons of the e.r.ps to high, medium and low probability words terminating sentences of medium constraint revealed larger N400 amplitudes to the less probable words [main effect of ending Cloze probability $F(2,26) = 12.94$, $P < 0.001$; ending $\times$ electrode $F(14,182) = 6.07$, $P < 0.001$].

Inspection of the e.r.ps to all types of terminal words (Fig. 1C) revealed a gradient of potential, with the greatest positivity following hi/hi words and a progressively larger negativity (N400) elicited by words of decreasing Cloze probability. This relationship between Cloze probability and N400 amplitude was also evidenced by the product-moment correlations between the two measures across the seven classes of word endings [for

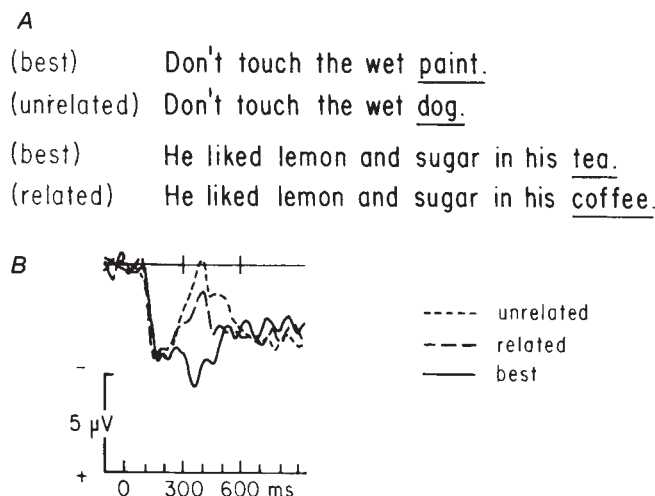


Fig. 2 **A**, Two examples of sentences with high contextual constraints completed by low Cloze probability words. Above each experimental sentence is the same sentence terminated by its 'best completion', which was or was not semantically related to the word that was actually presented in the sentence below. **B**, Grand average e.r.ps from Pz for the best completions (solid waveform), the semantically related (large dashed waveform) and the semantically unrelated (small dashed waveform) low Cloze probability words.

the grand average e.r.ps, Fz, 0.80; Cz, 0.90; Pz, 0.92; Oz, 0.92; L.AT, -0.82; R.AT, 0.94; L.PT, 0.88; R.PT, 0.97]. At the parietal site, this correlation calculated from individual subjects' data averaged 0.70 (0.21) and reached significance in seven subjects ($d.f. = 5$, $P < 0.05$). This correlation also held when the e.r.p. data were collapsed across the different levels of contextual constraint and re-averaged according to the Cloze probability of the final words [Fz, 0.87; Cz, 0.91; Pz, 0.94; Oz, 0.92; L.AT, -0.60; R.AT, 0.75; L.PT, 0.91; R.PT, 0.93]. The negative correlation observed at the left anterior temporal site could result from this scalp region being located on the opposite side of the dipole field of the N400 generator.

The systematic decline in the N400 amplitude as a function of increasing Cloze probability indicates that semantic incongruity is not a necessary condition for N400 elicitation. Instead, N400 amplitude appears to vary systematically as an inverse function of word expectancy, operationally defined here in terms of Cloze probability.

The influence of context on word recognition has been attributed to the automatic priming or activation of semantic networks, as well as to slower, attention-directed processes^{11–13}. Within such a framework, a sentence fragment primes (that is, activates for faster access) semantically related words whether or not they form acceptable sentence completions. In addition, sentence frames may result in the activation and retrieval of appropriate schemata^{14–16}. If the N400 reflects some aspect of this semantic activation, its amplitude should vary according to whether or not an unexpected terminal word is semantically related to the most expected ending of that sentence (that is to the 'best completion' of the activated schema).

We tested this prediction by reanalysing the e.r.ps to low Cloze probability words that completed highly constrained sentences, now segregating them according to whether or not the terminal word was related to the best completion (BC) of the sentence in which it occurred. The degree of semantic relatedness had been determined by asking a different group of 25 subjects to rate each word pair on a 5-point scale. The mean ratings were 4.03 (0.70) and 1.95 (0.73) for the related and unrelated word pairs, respectively. Sample sentences are shown in Fig. 2A. N400 amplitude was indeed sensitive to the semantic relationship between the eliciting word and the expected best completion (Fig. 2B), with larger N400s following words that were unrelated to the BC [main effect of semantic relatedness

$F(1,13) = 17.37$, $P < 0.001$; relatedness $\times$ electrode $F(7,97) = 3.14$, $P < 0.001$].

These results are in agreement with the hypothesis that the N400 component reflects the extent to which a word is semantically primed, rather than its being a specific response to contextual violations. If the N400 amplitude proves to be a valid index of semantic priming, it should become possible to investigate the timing and spread of activation within semantic networks and knowledge schemata and to identify automatic and attentional components of processing.

We thank J. C. Phillips and J. Hansen for their help. This work was supported by NSF grant BNS 80-05525 and Sloan Foundation grant B1900-35. M. K. is supported by Research Scientist Development Award USPHS-1K02MH0322.

Received 4 August; accepted 10 October 1983.

1. Picton, T. W. & Stuss, D. T. in *Biological Perspectives in Language* (eds Caplan, D. N., Lecours, A. R. & Smith, A. M.) (MIT Press, in the press).
2. Kutas, M. & Hillyard, S. A. *Science* **207**, 203-205 (1980).
3. Kutas, M. & Hillyard, S. A. *Mem. Cognit.* **11**(5), 539-550 (1983).
4. Sanquist, T. F., Rohrbaugh, J. W., Syndulko, K. & Lindsay, D. B. *Psychophysiology* **17**, 568-576 (1980).
5. Boddy, J. & Weinberg, H. *Biol. Psychol.* **12**, 43-61 (1981).
6. Fischler, I., Bloom, P. A., Childers, D. A., Roucos, S. E. & Perry, N. W. Jr *Psychophysiology* **20**, 400-409 (1983).
7. Polich, J., Vanasse, L. & Donchin, E. *Psychophysiology* **18**, 142 (1981).
8. Stuss, D. T., Sarazin, F., Leech, E. & Picton, T. W. in *Brain Information: Event-related Potentials Monogr.* 12 (eds Karrer, R., Cohen, J. & Tueting, P.) (New York Academy of Sciences, 1983).
9. Lesgold, A. M. & Perfetti, C. A. (eds) *Interactive Processes in Reading* (Erlbaum, Hillsdale, New Jersey, 1981).
10. Bloom, P. A. & Fischler, I. *Mem. Cognit.* **8**, 631-642 (1980).
11. Stanovich, K. E. in *Interactive Processes in Reading* (eds Lesgold, A. M. & Perfetti, C. A.) 241-267 (Erlbaum, Hillsdale, New Jersey, 1981).
12. Posner, M. I. & Snyder, C. R. R. in *Information Processing and Cognition: The Loyola Symposium* (ed. Solso, R.) (Erlbaum, Hillsdale, New Jersey, 1975).
13. Posner, M. I. & Snyder, C. R. R. in *Attention and Performance V* (eds Rabbitt, P. M. A. & Dornic, S.) 669-682 (Academic, New York, 1975).
14. Thorndyke, P. W. & Hayes-Roth, B. *Cognit. Psychol.* **11**, 82-106 (1979).
15. Kleiman, G. M. *Mem. Cognit.* **8**, 336-344 (1980).
16. Foss, D. J. *Cognit. Psychol.* **14**, 590-607 (1982).

Cell determination and regulation during development of neuroblasts and neurones in grasshopper embryo

Paul H. Taghert, Chris Q. Doe & Corey S. Goodman

Department of Biological Sciences, Stanford University, Stanford, California 94305, USA

The embryonic development of the central nervous system (CNS) involves the generation of an enormous diversity of cellular types arranged and interconnected in a remarkably precise pattern. In each hemisegment of the grasshopper embryo, the ectoderm generates a stereotyped pattern of 30 neuronal precursor cells, called neuroblasts (Fig. 1)¹. Each of these stem cells makes a stereotyped contribution of 6-100 progeny to the ~1,000 different neurones, each cell identifiable according to its unique morphology, physiology and biochemistry. What are the contributions of cell interactions and cell lineage to the generation of this diversity and specificity of identified neurones in the grasshopper CNS? Here we report on cell ablations with a laser microbeam at different stages of development. Our results suggest the importance of cell-cell interactions in the determination of ectodermal cells to become identified neuroblasts². However, once a neuroblast begins to divide, then cell lineage appears to play an important role in the determination of its stereotyped family of neuronal progeny³. Furthermore, cell-specific interactions continue to play an important role as neurones, according to their mitotic ancestry, recognize and interact with other differentiating neurones in their environment⁴⁻⁶.

Within the neuroepithelium in each segment of the grasshopper embryo, there arise 61 neuronal precursor cells, called neuroblasts (NBs), arranged in two symmetric plates of 30 NBs each and one median neuroblast (MNB) (Fig. 1)¹. We can

identify each NB by its position within the neuroepithelium; moreover, many NBs can now be uniquely identified by the highly stereotyped family of identified neurones they produce^{3,7-10}. Each NB is a stem cell which divides repeatedly to generate a chain of ganglion mother cells. Each ganglion mother cell then divides once more to generate two ganglion cells in a chain of cell doublets that then differentiate into neurones. The 30 NBs in each hemisegment largely generate the ~1,000 neurones in each thoracic hemiganglion and the ~250 neurones in each abdominal hemiganglion; most if not all of these neurones can be uniquely identified according to their characteristic morphology, physiology and biochemistry. The number of neuronal progeny appears to be specific for individual NBs; for example, NB 7-3 produces 6 progeny¹⁰ whereas the MNB produces ~100 progeny⁶⁻⁸. Furthermore, the identities of the progeny appear to be specific for individual NBs; the fate of individual neurones is highly correlated with their specific positions in the family trees of particular NBs⁶⁻¹⁰.

The progeny of NBs form coherent families of neurones that are inserted into the developing CNS in characteristic positions. The determination of an individual neurone might be due to its position in the NB family tree, or alternatively to its position in the developing ganglion and thus its reproducible interactions with its neighbours. Selective ablation of identified cells with a laser microbeam can help to distinguish between these alternatives. For example, NB 7-3 gives rise to six progeny (Fig. 1). In the abdominal segments (A4-A6), the first two NB 7-3 progeny are called S1 and S2; following their differentiation they can be specifically stained with an anti-serotonin antibody (Fig. 2). They are the only prominent serotonin-immunoreactive neurones in these segments. Furthermore, S1 and S2 can also be individually identified by their distinctive morphologies after intracellular dye injections (Fig. 1)¹. We ablated NB 7-3 *in ovo* just before it begins its first cell division to test if some other NB could produce neurones with either the biochemical or morphological phenotypes of S1 and S2 (Fig. 2). We also ablated other NBs at the same stage to test if interactions with neighbouring NBs or their progeny were necessary for NB 7-3 to produce neurones S1 and S2.

NBs were identified in a 30%¹¹ embryo through the dechorionated eggcase under brightfield optics, and ablated with a pumped-dye laser (Fig. 2). Experimental and control embryos were examined at three different stages with three different techniques: at 33% with Nomarski optics to look for NB 7-3; at 50% with intracellular dye injections of Lucifer Yellow to look for cells with the distinctive morphology of neurones S1 and S2; and at 70% with the anti-serotonin antibody to look for cells with the distinctive serotonin-immunoreactivity of neurones S1 and S2.

When NB 7-3 was ablated before it began its first cell division and the embryos were assayed at 33% (the following day) for the presence or absence of the NB, we found that in approximately 80% of the cases ($n > 20$) no regulation took place. When regulation does occur, a new NB 7-3 is present as are all neighbouring NBs. When embryos were assayed at 70%, in 7 of 10 cases no serotonin-immunoreactive neurones were present on the experimental side of the ganglion (Fig. 2). In these embryos, the homologous S1 and S2 neurones from the contralateral NB 7-3 were always present and had the normal morphological and biochemical features. These observations suggest that when a NB is ablated before it begins its first cell division, regulation does not usually take place. In addition, progeny of neighbouring NBs do not differentiate with either the morphological or biochemical properties of neurones S1 or S2. The partial (20-30%) regulation we observe following NB ablation is likely to result from the recruitment of a neighbouring ectodermal cell, as described in the *in vitro* experiments below. In three other cases in which NB 7-3 was ablated, exhaustive intracellular dye injections at 50% of every other cell body in the region where S1 and S2 normally appear revealed no neurones with the distinctive morphology of S1 and S2 on the experimental side; S1 and S2 were easily found on the control side, and did

Fig. 1 A, Schematic diagram of a 30% grasshopper embryo¹¹ (hatching occurs at 100%) showing the segmental structure of the ectoderm. The middle of the ectoderm is a longitudinal strip of neuroepithelium which gives rise to the CNS. The location of the neuronal precursor cells for a single segment (A4) are drawn in black. B, Pattern of NBs for a single segment (A4) includes two plates of 30 NBs arranged in a precise pattern of seven rows, and one MNB making a total of 61 NBs. C, Cell lineage of the two identified serotonin immunoreactive neurones, S1 and S2, from NB 7-3 in segment A4. Cells S1 and S2 arise by an invariant cell lineage from the first division of NB 7-3. NBs generate neuronal progeny by a series of asymmetric cell divisions that produce small, ganglion mother cells. Ganglion mother cells then each divide once more symmetrically to produce two cells that will both differentiate into neurones. The cell lineage is represented here with a ganglion mother cell closest to the NB and older neuronal progeny lying progressively farther away from the NB. D, The cell-specific morphologies of neurones S1 (upper) and S2 (lower) in the A4 ganglion at 70% of embryonic development, as revealed by intracellular injection of Lucifer Yellow followed by HRP immunocytochemistry with an anti-Lucifer Yellow antibody¹⁸.

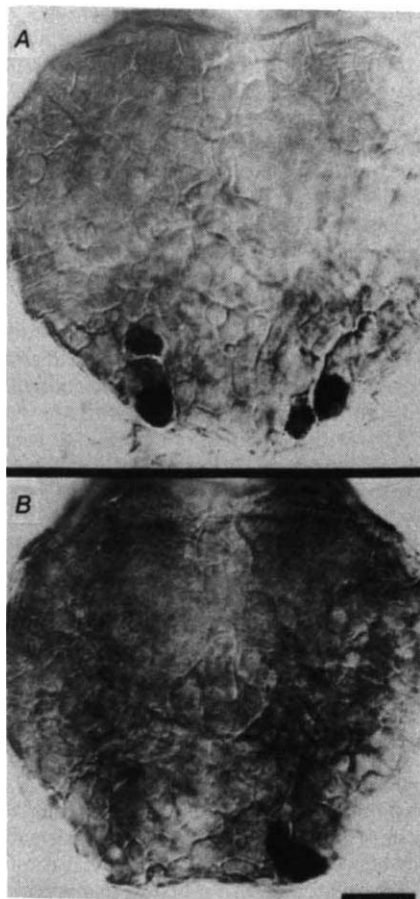
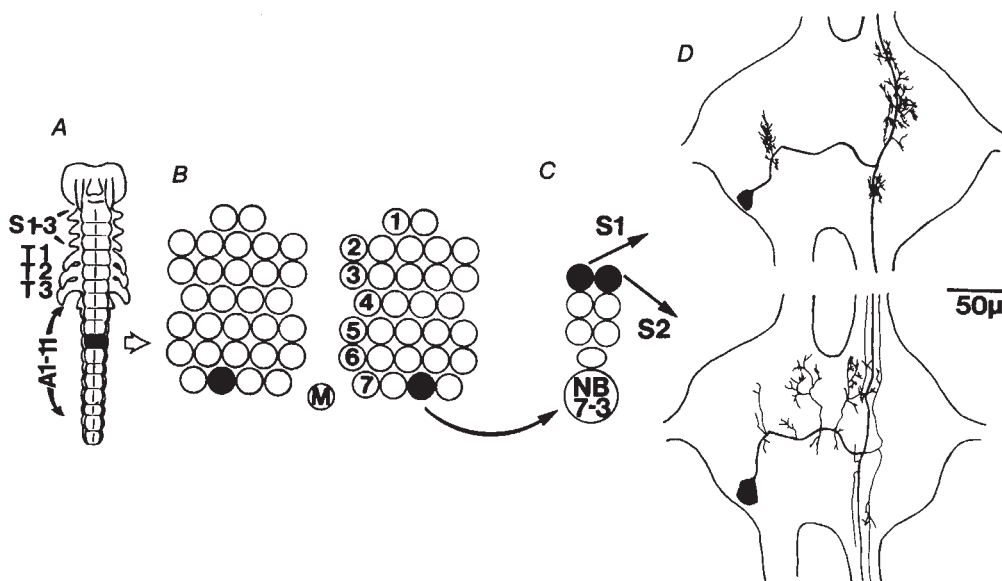


Fig. 2 Laser ablation of NB 7-3 *in ovo* before it began its cycle of cell divisions (30% of development) led to the absence of its identified progeny, neurones S1 and S2, in segment A4. A, Photomicrograph of an unoperated control ganglion stained in wholemount with the anti-serotonin antiserum at 70% of embryonic development. Note the bilaterally symmetrical immunoreactive cells, S1 and S2. B, Photomicrograph of a similarly stained ganglion in which NB 7-3 had been laser ablated earlier in development. Note the absence of immunoreactive cells on the operated side. Scale bar, 30 µm.

Methods: 30% Embryos were dechorionated in bleach, rinsed in saline and gently flattened with a coverslip to improve visibility of the NBs. NBs were identified through the dechorionated eggcase under brightfield optics with a Zeiss compound microscope using a Leitz 50× water immersion lens. Laser ablations were performed with a Chromaser pumped-dye laser system (Phase-R Corporation) using Coumarin 504 dye. Target cells were selected with a colinear helium-neon laser; 5 to 10 shots at ~250 mJ per shot over 5 to 10 min were used to kill NBs. Condensation of chromatin and subsequent loss of cellular integrity were seen within this time. Eggs were discarded given any sign of damage to neighbouring, non-target cells. Eggs with specific cellular ablations were removed from the slide and incubated. For serotonin immunostaining, the isolated embryonic CNS was fixed in 2% paraformaldehyde in Millonig's buffer and then processed according to the peroxidase-anti-peroxidase method of Sternberger¹⁹. Tissues were preincubated in 10% normal goat serum, 2% bovine serum albumin (BSA), and 1% Triton X-100 in phosphate-buffered saline (PBS) for 1 h. Tissues were then incubated in anti-serotonin antibody (Immunonuclear) at a 1:500 dilution, 1% normal goat serum, 1% bovine serum albumin and 1% Triton X-100 in PBS overnight at 4 °C. Following incubations in linking antibody and the peroxidase-anti-peroxidase antibody complex for 1 h each at room temperature, the tissues were developed for the horseradish peroxidase reaction product, cleared and examined.

not dye couple (as they normally do) to their homologues on the experimental side.

Not only does regulation not usually take place, but NB 7-3 appears determined to produce a particular family of neurones. Ablation of neighbouring NBs [for example, NB 6-5 ($n=20$) which produces a small family, and NBs 7-2 and 7-4 ($n=4$) which produce large families] did not appear to alter the morphological or biochemical differentiation of these first two NB 7-3 progeny as revealed with the anti-serotonin antibody. Furthermore, there was no apparent increase in the size of the

NB 7-3 family to compensate for the loss of neighbouring families. Thus, cell lineage does appear to determine some aspects of cell fate in the insect CNS; NBs appear determined to produce particular families of neurones which differentiate with particular phenotypes.

But in what fashion does the precise pattern of identified NBs arise? Do they arise as a consequence of their cell lineage, as occurs for neuronal precursor cells in the nematode¹² and leech¹³, or do they arise by cell-cell interactions, as occurs for the sensory neurone precursor cells in the epidermis of insects¹⁴?

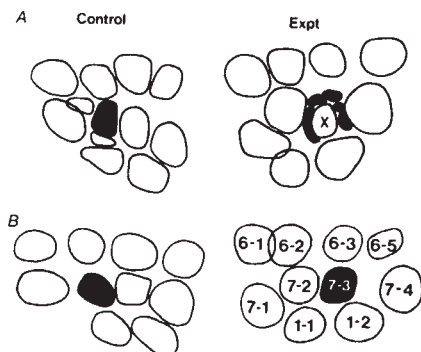


Fig. 3 *In vitro* laser ablation of ectodermal cell enlarging to become NB 7-3. Camera lucida drawings of NB rows 6 and 7 in segment T1, and NB row 1 in adjacent segment T2. **A**, A 26% embryo was removed from its eggcase under sterile conditions and viewed with a Zeiss compound microscope using Nomarski optics and a Leitz $\times 50$ water immersion lens for cell identifications. At the future position of NB 7-3 on the control side, the enlarging ectodermal cell (shaded) was allowed to develop normally, whereas its contralateral homologue (marked by X) was killed with three shots of the laser microbeam at 200 mJ per shot. This ablation left 5 smaller ectodermal cells (shaded) in the presumptive NB 7-3 position. The embryo was then cultured for 24 h in an RPMI 1640 culture medium containing glutamine (0.29 g l^{-1}), sodium pyruvate (0.11 g l^{-1}), penicillin (50 IU l^{-1}), streptomycin ($50 \mu\text{g ml}^{-1}$), sodium bicarbonate (2 g l^{-1}), horse serum (10%), glycine (6.0 g l^{-1}), bovine insulin (7.6 mg l^{-1}), 20-hydroxyecdysone ($15 \mu\text{g ml}^{-1}$), and juvenile hormone I ($0.5 \mu\text{g ml}^{-1}$)⁵. Cultures were maintained in a moist incubator in a 5% CO_2 atmosphere at 29°C . **B**, Camera lucida drawing of the same NB rows as in **A**, but after 24 h in culture. NB 7-3 (shaded) developed on both the control and experimental sides.

In insect embryos, gastrulation leads to an inner layer of mesodermal cells and an outer layer of ectodermal cells. It is from the seemingly uniform sheet of ectodermal cells between the midline and limb buds that the stereotyped pattern of NBs develops. Not all NBs form simultaneously. Initially, NBs arise from ectodermal cells that enlarge and delaminate inwards. Later, NBs arise from ectodermal cells that have delaminated to form clusters of small cells between the older NBs (Fig. 3). Ultimately, only one cell within each cluster enlarges and becomes a NB; the remaining cells either die or form ensheathing cells around the NB and sometimes its progeny. There are between three and six ectodermal cells in the region between NBs 7-2 and 7-4 from which NB 7-3 arises. To distinguish between cell lineage and cell-cell interactions in NB determination, we used a laser microbeam *in vitro* to ablate individual cells within the cluster of ectodermal cells from which NB 7-3 normally arises (Fig. 3)².

When we ablated the entire cluster of small cells, one of which would normally form NB 7-3, the NB did not appear ($n=2$). When we ablated only a portion of the cluster, however, NB 7-3 formed. In particular, when we killed only the cell that was enlarging to become NB 7-3, regulation did take place ($n=6$); one of the remaining cells became NB 7-3 (Fig. 3)². The new NB 7-3 generated progeny by stereotyped asymmetric divisions, but embryos cultured *in vitro* did not develop far enough to determine the morphological or biochemical identity of these progeny. Apparently NB 7-3 can form from any one of the cells within the cluster, even after one cell has begun enlarging. Thus, cell-cell interactions within the cluster appear to determine which cell becomes the NB.

Our results demonstrate that at early stages no one cell is uniquely determined to be NB 7-3, but rather NB 7-3 develops from a restricted group of cells that share a common potential. Cell-cell interactions appear to assure that only one cell acquires the NB 7-3 fate. We do not know if the specification of individual NBs occurs in one step, or alternatively if an ectodermal cell first acquires a general NB determination and subsequently acquires a unique NB fate. By the time NB 7-3 begins its first

cell division, however, it appears uniquely committed to generate a specific family of neurones; no other cell, including ectodermal cells adjacent to the ablated NB, can take its place. The NB 7-3 progeny cannot be replaced by progeny of neighbouring NBs. The final differentiation of each individual neurone depends on its mitotic ancestry and its later cell-specific interactions with other neurones in its environment⁴⁻⁶.

These results suggest that the mechanisms giving rise to precursors for motoneurons and interneurons in the CNS are similar to those giving rise to precursors for sensory neurones¹⁴ in the periphery of insects. Both types of precursors are derived from the ectoderm. Both undergo stereotyped patterns of cell divisions, the peripheral precursors typically generating four cells, one of which becomes a neurone, while the central precursors each generate 6-100 cells, all of which become neurones. The determination of both types of precursors appears to be based on position and not lineage, and involves cell-cell interactions that lead to an inhibition of surrounding cells from assuming similar fates. The differences are simply in the precision and complexity of the pattern. The central precursors develop in a highly stereotyped array while the peripheral precursors are more sparsely and often less precisely arranged. Moreover, while the peripheral precursors produce only a few different types of lineages (bristle, campaniform sensillum and so on), each central precursor appears to be uniquely determined to produce a specific lineage of neurones, each NB differing from the others in both the number and phenotypes of its neuronal progeny. Thus, the pattern formation in the CNS may involve a more complex form of the same mechanisms underlying pattern formation in the epidermis. Whether making the CNS involves more complex cell-cell interactions, or subdivisions of the ectodermal compartments¹⁵⁻¹⁷ into smaller units than those required to generate the epidermis, remains to be tested.

We thank John Thomas for criticism of the manuscript. This work was supported by a NIH postdoctoral fellowship to P.H.T., NSF predoctoral fellowship to C.Q.D., and a NIH grant and McKnight Scholars Award to C.S.G.

Received 8 August; accepted 3 October 1983.

1. Bate, C. M. *J. Embryol. exp. Morph.* **35**, 107-123 (1976).
2. Doe, C. Q. & Goodman, C. S. *Soc. Neurosci.* **9**, 899 (1983).
3. Taghert, P. H. & Goodman, C. S. *J. Neurosci.* (in the press).
4. Raper, J. A., Bastiani, M. J. & Goodman, C. S. *Soc. Neurosci.* 1044 (1983).
5. Raper, J. A., Bastiani, M. J. & Goodman, C. S. *Cold Spring Harb. Symp. quant. Biol.* **48**, (in the press).
6. Bastiani, M. J. & Goodman, C. S. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
7. Goodman, C. S. & Spitzer, N. C. *Nature* **280**, 208-214 (1979).
8. Goodman, C. S. & Bate, M. *Trends Neurosci.* **4**, 163-169 (1981).
9. Goodman, C. S. in *Neuronal Development* (ed. Spitzer, N. C.) (Plenum, New York, 1982).
10. Raper, J. A., Bastiani, M. & Goodman, C. S. *J. Neurosci.* **3**, 20-30 (1983).
11. Bentley, D., Keshishian, H., Shankland, M. & Toroian-Raymond, A. *J. Embryol. exp. Morph.* **54**, 47-74 (1979).
12. Sulston, J. E. & Horvitz, H. R. *Dev. Biol.* **56**, 110-156 (1977).
13. Stent, G. S., Weisblat, D. A., Blair, S. S. & Zackson, S. L. in *Neuronal Development* (ed. Spitzer, N. C.) (Plenum, New York, 1982).
14. Lawrence, P. A. in *Developmental Systems in Insects* Vol. 2 (eds Counce, S. J. & Waddington, C. H.) (Academic, New York, 1973).
15. Garcia-Bellido, A., Ripoll, P. & Morata, G. *Nature* **245**, 251-253 (1973).
16. Crick, F. H. C. & Lawrence, P. A. *Science* **189**, 340-347 (1975).
17. Lawrence, P. A. *Cell* **26**, 3-10 (1981).
18. Taghert, P. H., Bastiani, M. J., Ho, R. H. & Goodman, C. S. *Dev. Biol.* **94**, 391-399 (1982).
19. Sternberger, L. A. *Immunocytochemistry* (Wiley, New York, 1978).

Prostaglandin I_2 is formed *in vivo* in man after dietary eicosapentaenoic acid

Sven Fischer & Peter C. Weber

Medizinische Klinik Innenstadt der Universität München, Ziemssenstraße 1, 8000 München 2, FRG

Greenland Eskimos who live on a traditional marine diet rich in long chain ω -3-polyunsaturated fatty acids have a low incidence of cardiovascular disease and myocardial infarction¹⁻³. In their plasma and platelet lipids, arachidonic acid, the precursor of dienoic prostanoids, is partly replaced by eicosapentaenoic acid ($\text{C}_{20:5}$, ω 3; EPA), the precursor of trienoic prostanoids^{4,5}. Studies with an Eskimo diet⁶ or a

Western diet supplemented with sea fish or fish oil rich in EPA⁷⁻¹⁴ resulted in an 'Eskimo-like' pattern of plasma and platelet lipids. Moreover, less reactive platelets, a reduced *ex vivo* formation of proaggregatory thromboxane A₂ and a blunted circulatory response to pressor hormones were reported. These favourable functional effects may be induced by a shift of prostanoid formation from the dienoic to the trienoic series^{14,15}. We show here that the major urinary metabolite of endogenous prostaglandin I₃ is present in subjects that have ingested either cod liver oil (~4 g EPA per day) or mackerel (~10-15 g EPA per day). Our studies provide the first direct evidence for *in vivo* formation of prostaglandin I₃ in man.

In vitro experiments have shown that EPA-derived thromboxane A₃ (TXA₃), in contrast to arachidonic acid (C20:4, ω 6; AA)-derived thromboxane A₂ (TXA₂)¹⁶, does not^{17,18} or only weakly¹⁹ aggregate platelets, whereas EPA-derived prostaglandin I₃ (PGI₃)^{15,18,20} is as antiaggregatory as prostaglandin I₂ (PGI₂)²¹. Consequently, a shift of prostanoid formation from the dienoic to the trienoic series may change the TXA/PGI balance²² in a favourable way¹⁵ (Table 1).

Previous studies have demonstrated formation of PGI₃ from exogenous EPA or endoperoxide PGH₃ in certain *in vitro* systems including rat arterial tissue¹⁵, bovine and porcine aortic microsomes^{18,23} and human umbilical arteries²⁴. Tissue formation of trienoic prostaglandins has been found in bovine lung²⁵ and kidney homogenates from rats fed with fish oil²⁶. The physiological significance of these findings has been uncertain because comparative studies using purified cyclooxygenase from sheep vesicular glands and AA and EPA as substrates²⁷ indicate that a high peroxide level, presumably not present *in vivo*, might be necessary for an appreciable formation of trienoic prostaglandins. Moreover, in other EPA feeding studies in the rat, tissue formation of PGI₃ was not detectable²⁸.

We addressed three major questions in this study: (1) Is there a substantial formation of PGI₃ *in vivo* in healthy man after a diet rich in EPA and is the intake of EPA associated with a decrease in PGI₂ formation? An analogous alteration has been found in platelet studies where EPA reduced formation of TXA₂ (refs 8, 14, 18). (2) Is there a quantitative relationship between the amount of EPA ingested, the induced changes of fatty acids in plasma phospholipids and the formation of PGI₂/PGI₃? (3) Are there changes in platelet aggregability, formation of TXA₂ and bleeding time after dietary EPA?

In our study we extracted, before and after dietary supply of EPA, dinor metabolites of PGI according to a highly selective method²⁹ from urine collected over 24-h periods. We used Δ 17-2,3-dinor-6-keto-prostaglandin F_{1 α} (PGI₃-M) as an index of total body production of PGI₃ in analogy to 2,3-dinor-6-keto-prostaglandin F_{1 α} (PGI₂-M), which has been shown to be the major urinary metabolite of PGI₂ (refs 29, 30). A prerequisite for identification and quantification of PGI₃-M by mass spectrometry was its separation from PGI₂-M. This was achieved by using a capillary gas chromatography column with high resolving power. A partial mass spectrum of endogenous PGI₃-M extracted from urine after dietary intake of EPA is shown in Fig. 1. For comparison a partial mass spectrum of PGI₂-M simultaneously extracted from the same urine sample is presented below. The characteristically different fragmentation of PGI₃-M is due to the additional Δ 17-double bond in this molecule. Fragments m/z 530 (M^+ -69(71)), 440 (M^+ -90-69(71)) and 350 (M^+ -2 $\times$ 90-69(71)) are common to both metabolites, but are more intense in PGI₃-M, due to preferred allylic cleavage of the methyl side chain (C₁₆-C₂₀). Fragments m/z 584 (M^+ -15), 568 (M^+ -31), 478 (M^+ -90-31) and 388 (M^+ -2 $\times$ 90-31) are characteristic of PGI₃-M, but are less intense than fragments 586 (M^+ -15), 570 (M^+ -31), 480 (M^+ -90-31) and 390 (M^+ -2 $\times$ 90-31) which are characteristic of PGI₂-M. Some unidentified peaks in the spectra result from biological contaminating compounds in the urinary lipid extract which co-elute from the capillary column.

The multiple ion-selection tracings shown in Fig. 2 illustrate that PGI₃-M elutes from the capillary column about 35 s after

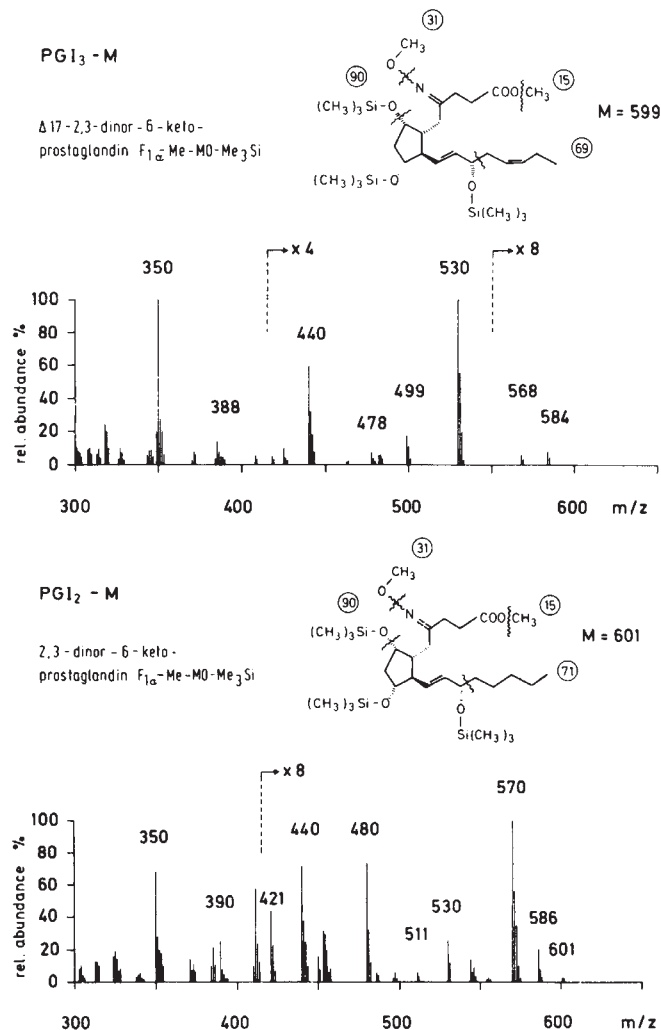


Fig. 1 Partial mass spectra (Me, MO, Me₃Si derivatives) of endogenous Δ 17-2,3-dinor-6-keto-PGF_{1 α} and 2,3-dinor-6-keto-PGF_{1 α} , the main urinary metabolites of PGI₃ and PGI₂, respectively. 800 ml of urine, collected after daily ingestion of 750 g mackerel for 3 days, were extracted according to the method of Falardeau *et al.*²⁹. This method extracts the dinor metabolites of PGI₂ and PGI₃ highly selectively due to their ability to form γ -lactones in acidic solution. Extracts were methylated with ethereal diazomethane and further purified on silicagel TLC using the solvent system ethylacetate/isooctane/acetic acid/water = 110/50/20/100 (organic phase). The R_F value of authentic methylated 2,3-dinor-6-keto-PGF_{1 α} was 0.24. The Me, MO, Me₃Si derivatives for GC-MS were prepared as described previously²⁹. Prominent ions in the range above m/z 300 for Δ 17-2,3-dinor-6-keto-PGF_{1 α} -Me,MO,Me₃Si are: 584 (M^+ -15), 568 (M^+ -31), 530 (M^+ -69), 499 (M^+ -69-31), 478 (M^+ -90-31), 440 (M^+ -90-69), 388 (M^+ -2 $\times$ 90-31) and 350 (M^+ -2 $\times$ 90-69); for 2,3-dinor-6-keto-PGF_{1 α} -Me,MO,Me₃Si: 601 (M^+), 586 (M^+ -15), 570 (M^+ -31), 530 (M^+ -71), 511 (M^+ -90), 480 (M^+ -90-31), 440 (M^+ -90-71), 421 (M^+ -2 $\times$ 90), 390 (M^+ -2 $\times$ 90-31) and 350 (M^+ -2 $\times$ 90-71). The GC-MS system was a Finnigan MAT 44S equipped with a 50-m fused silica WCOT capillary column (Carbowax CP 51, 0.23 mm internal diameter; Chrompack), separating prostaglandins of the two and three series. Operation conditions of the GC-MS system were: injection port 280 °C; interface 270 °C; ion source 200 °C, electron impact energy 80 eV; current of emission 0.9 mA; electron multiplier voltage 1.8 kV.

PGI₂-M and is detectable only after dietary intake of EPA. Chemical reactions were performed to confirm the structure of PGI₃-M. Derivatization with methoxy-²H₃-amine hydrochloride resulted in an increase of the molecular weight of PGI₂-M and PGI₃-M by 3 mass units and established the presence of one keto group in both molecules. Identity of PGI₃-M was confirmed by catalytic hydrogenation of the urinary extract.

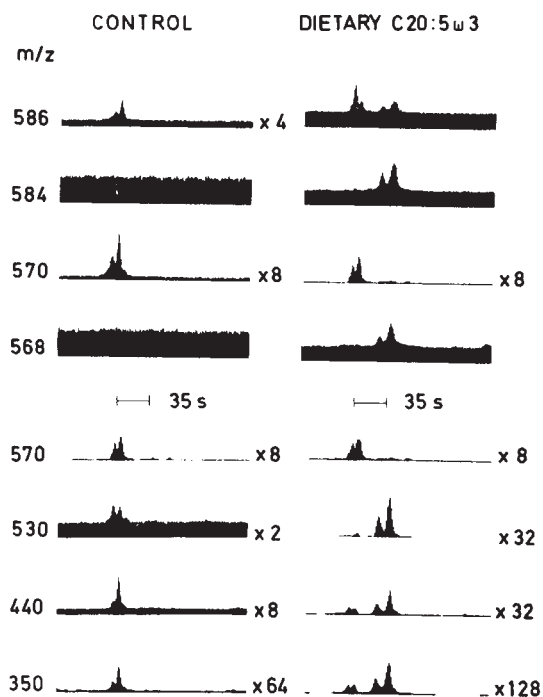


Fig. 2 Multiple ion-selection tracings (Me,MO,Me₃Si derivatives) of PGI₂-M extracted from a control urine (left) and of PGI₂-M and PGI₃-M extracted from a urine collected after daily ingestion of 750 g mackerel for 3 days (right). Qualitatively similar tracings were found after ingestion of cod liver oil. Double peaks are due to *syn*, *anti*-isomers formed by the methoxime group of the derivatives. A time axis indicates the elution time between both metabolites. Sensitivity of the tracings is lowered by the factors indicated and is adjusted to the highest occurring peak. Fragments 586 (M⁺-15) and 570 (M⁺-31) are specific for PGI₂-M, whereas analogous fragments 584 (M⁺-15) and 568 (M⁺-31) are specific for PGI₃-M. Fragments *m/z* 530 (M⁺-69(71)), 440 (M⁺-90-69(71)) and 350 (M⁺-2×90-69(71)) are common to both metabolites. Deuterated internal standard (2,3-dinor-6-keto-19,19',20,20'-²H-PGF_{1α}) for quantification of PGI₂-M has not been added to this sample.

Table 1 Physiological properties of the conversion products of arachidonic acid and eicosapentaenoic acid in vessel wall and platelets

Precursor acid	Arachidonic acid	Eicosapentaenoic acid
Vessel wall	PGI ₂ Antiaggregatory	PGI ₃ Antiaggregatory
Platelets	TXA ₂ Proaggregatory	TXA ₃ Weakly proaggregatory

Table 2 Polyunsaturated fatty acid metabolism after intake of cod liver oil or mackerel

	Control <i>n</i> = 6	Cod liver oil Day 25 <i>n</i> = 6	Control <i>n</i> = 3	Mackerel Day 1 <i>n</i> = 3	Mackerel Day 3 <i>n</i> = 3
Rel. %					
C20:5ω3	0.7 ± 0.3	6.5 ± 1.2*	1.1 ± 0.7	7.5 ± 1.8*	12.2 ± 4.7*
C22:6ω3	4.0 ± 0.8	8.7 ± 1.3*	4.1 ± 1.5	5.0 ± 1.2	6.2 ± 0.8
C20:4ω6	9.8 ± 1.9	7.9 ± 1.2	8.7 ± 2.6	7.2 ± 2.4	6.8 ± 1.1
C18:2ω6	23.2 ± 2.1	14.1 ± 2.4*	21.1 ± 6.0	7.9 ± 1.4*	5.6 ± 1.9*
ng per 24 h					
PGI ₂ -M	146 ± 39	162 ± 52	121 ± 51	192 ± 27†	236 ± 32†
PGI ₃ -M	Not detectable	83 ± 25	Not detectable	114 ± 32	134 ± 48

Patterns of polyunsaturated fatty acids (% of total fatty acids, mean ± s.d.) in plasma phospholipids and urinary excretion of PGI₂-M and PGI₃-M (ng per 24 h, mean ± s.d.) after intake of cod liver oil (40 ml ± 4 g EPA daily for 24 days; unchanged Western diet) or mackerel (750 g ± 10–15 g EPA daily for 3 days; no intake of other proteins or lipids). Fatty acids in plasma phospholipids were quantified by gas chromatography as described previously⁸. PGI₂-M and PGI₃-M were extracted simultaneously from urine according to the method of Falardeau *et al.*²⁹. PGI₂-M was quantified by GC-MS using 2,3-dinor-6-keto-19,19',20,20'-²H-PGF_{1α} as the deuterated internal standard and monitoring the ion pair *m/z* 570/574 (M⁺-31). The analysis was based on a standard curve which was shown to be linear in the observed range³⁴. Excretion of PGI₃-M could not be quantified by GC-MS via a direct method due to lack of deuterated PGI₃-M or authentic unlabelled material. Quantification using deuterated PGI₂-M was also not possible because fragments suitable for multiple ion-selection represent different percentages of the total ion current of each molecule. Therefore, PGI₃-M excretion was quantified by an indirect method assessing at first the total excretion of PGI₂-M and PGI₃-M. This was achieved after catalytic hydrogenation of the urinary extract containing deuterated PGI₂-M as internal standard by monitoring the ion pair *m/z* 572/576 (M⁺-31) of hydrogenated PGI_{2,3}-M and hydrogenated deuterium labelled internal standard (Me, MO, Me₃Si derivatives) which in their hydrogenated form represent compounds with identical fragmentation patterns. Urinary excretion of PGI₃-M was then calculated by subtracting the previously measured amount of PGI₂-M from the total amount of hydrogenated PGI_{2,3}-M. The method for quantification of urinary PGI₃-M was validated by a recovery experiment. Increasing amounts of PGI₂-M and a constant amount of deuterated PGI₂-M were added to a urine voided after mackerel ingestion. After hydrogenation of the urinary extract the added PGI₂-M was recovered quantitatively by GC-MS. Cod liver oil was manufactured by Møller A/S. All solvents and silicagel TLC plates were from Merck-Darmstadt. The 2,3-dinor-6-keto-PGF_{1α}-lactone and the 2,3-dinor-6-keto-19,19',20,20'-²H-PGF_{1α} were a gift from Dr J. Pike (Upjohn). Methoxy-²H₃-amine hydrochloride was prepared according to ref. 35. Catalytic hydrogenation was performed in the presence of Rh/charcoal in ethyl acetate.

* *P* < 0.01; † *P* < 0.05 compared with control values.

A single new compound was formed from PGI₂-M and PGI₃-M with a molecular weight of *m/z* 603, corresponding to an uptake of 1 and 2 equivalents of hydrogen, respectively. This substance co-chromatographed on capillary gas chromatography with authentic hydrogenated PGI₂-M and exhibited identical fragments with analogous intensities.

Table 2 summarizes the analyses of ω3 and ω6 polyunsaturated plasma phospholipid fatty acids and the urinary excretion rates of PGI₂-M and PGI₃-M after two separate studies with different amounts of dietary EPA. Polyunsaturated plasma phospholipid fatty acids indicated a characteristic change. Whereas EPA increased drastically during both dietary regimens, C22:6,

ω3 was elevated significantly only after ingestion of cod liver oil due to the higher content of this fatty acid in cod liver oil as compared with mackerel. AA only decreased slightly after cod liver oil or mackerel ingestion, whereas C18:2, ω6 was strongly reduced after these dietary regimens. The more pronounced decrease of C18:2, ω6 in plasma phospholipids during the mackerel diet probably resulted from the complete substitution of animal and vegetable ω6-fatty acids by ω3-fatty acids of the mackerel. Excretion of PGI₃-M was not detectable in control conditions, but was clearly measurable after intake of cod liver oil and increased with increasing EPA incorporation into plasma phospholipids, reaching levels one-half of those

measured for PGI₂-M. The content of EPA in plasma phospholipids and the urinary excretion rates of PGI₃-M were elevated within 1 day of mackerel ingestion and increased only slightly thereafter. This may indicate that endothelial cells rapidly equilibrate with the plasma lipids and thereby transform EPA to PGI₃, as has been shown for PGI₂ from AA³¹. Our results also show that dietary EPA does not inhibit *in vivo* formation of PGI₂ from AA, confirming similar results from *in vitro* studies using human umbilical vasculature and AA and EPA as substrates³². Excretion of PGI₂-M was unchanged (cod liver oil) or even increased (mackerel) under the influence of dietary EPA. The reason for the increase of PGI₂-M excretion after mackerel ingestion is unclear. According to observations in endothelial cell cultures³³, the drastic decline of C18:2, ω 6 in plasma phospholipids after mackerel ingestion may contribute to an increased prostacyclin synthetase activity in endothelial cells.

We have previously shown that platelet aggregation induced by low concentrations of collagen and associated formation of immunoreactive TXB₂ are reduced after a diet of mackerel in healthy subjects⁸. In the present study also, we found that ingestion of cod liver oil (CLO) for 25 days resulted in a reduced platelet aggregation (measured as % light transmission) in platelet rich plasma on stimulation with low doses (0.75 μ g ml⁻¹) of collagen (control, 70 $\pm$ 5% (s.d.); CLO, 53 $\pm$ 15%; P < 0.01). Concomitantly, a reduced formation of immunoreactive TXB₂ in collagen stimulated platelets (control, 40 $\pm$ 13 ng ml⁻¹; CLO, 15 $\pm$ 2.5 ng ml⁻¹; P < 0.01) and an increase of bleeding time (control, 104 $\pm$ 34 s; CLO, 145 $\pm$ 52 s; P < 0.01) were observed. Furthermore, in recent experiments³⁶ we have traced by combined gas chromatography-mass spectrometry (GC-MS) *ex vivo* formation of weakly proaggregatory TXA₃ in human platelets after supplementation of the diet with cod liver oil. Thus, the reduced platelet aggregability, increased bleeding time and reduced blood pressure response to pressor hormones¹⁴ after dietary EPA may be due to formation of PGI₃ and TXA₃ associated with a reduced synthesis of TXA₂ and an unchanged or even increased formation of PGI₂.

Our findings with a diet rich in EPA prove that it is possible in man to change the spectrum of biologically highly active prostanoids by nutritional means and alter it in a favourable direction. This may be a mechanism by which EPA-enriched diets induce a less thrombogenic state and hence may be beneficial for the prevention of atherothrombotic disorders.

We thank Drs B. Scherer, R. Lorenz and W. Meister for cooperation, Dr C. von Schacky for performing parts of the fatty acid analyses and U. Katzner for technical assistance. This study was supported by Deutsche Forschungsgemeinschaft.

Received 27 June; accepted 10 November 1983.

1. *Lancet* i, 1139-1141 (1983).
2. Bertelsen, A. *Medd. Grönland* 117, 1-83 (1935).
3. *The State of Health in Greenland*, Annual reports 1973, 1974, 1975, 1976 (Godthab, Denmark, 1978).
4. Dyerberg, J., Bang, H. O. & Hjørne, N. *Am. J. clin. Nutr.* 28, 958-966 (1975).
5. Dyerberg, J. & Bang, H. O. *Lancet* ii, 433-435 (1979).
6. Sinclair, H. M. in *Drugs Affecting Lipid Metabolism* (eds Fumagalli, R., Kritchevsky, D. & Paoletti, R.) 363-370 (Elsevier, Amsterdam, 1980).
7. v. Lossnocy, T. O., Ruiters, A., Bronsgeest-Schoute, H. C., van Gent, O. M. & Hermus, R. J. *J. Am. J. clin. Nutr.* 31, 1340-1346 (1978).
8. Siess, W. *et al. Lancet* i, 441-444 (1980).
9. Sanders, T. A. B., Vickers, M. & Haines, A. P. *Clin. Sci.* 61, 317-324 (1981).
10. Goodnight, S. H., Harris, W. S. & Connor, W. E. *Blood* 58, 880-885 (1981).
11. Thorngren, M. & Gustafson, A. *Lancet* ii, 1190-1193 (1981).
12. Brox, J. H., Killie, J. E., Gunnes, S. & Nordøy, A. *Thromb. Haemostasis* 46, 604-611 (1981).
13. Hirai, A. *et al. Thromb. Res.* 28, 285-298 (1982).
14. Lorenz, R., Spengler, U., Fischer, S., Duhm, J. & Weber, P. C. *Circulation* 67, 504-511 (1983).
15. Dyerberg, J., Bang, H. O., Stoffensen, E., Moncada, S. & Vane, J. R. *Lancet* ii, 117-119 (1978).
16. Hamberg, M., Svensson, J. & Samuelsson, B. *Proc. natn. Acad. Sci. U.S.A.* 72, 2994-2998 (1975).
17. Needleman, P., Minkes, M. & Raz, A. *Science* 193, 163-165 (1976).
18. Needleman, P., Raz, A., Minkes, M. S., Ferrandelli, J. A. & Sprecher, H. *Proc. natn. Acad. Sci. U.S.A.* 76, 944-948 (1979).
19. Gryglewski, R. J. *et al. Prostaglandins* 18, 453-478 (1979).
20. Nidy, E. G. & Johnson, R. A. *Tetrahedron Lett.* 27, 2375-2378 (1978).
21. Moncada, S., Gryglewski, R., Bunting, S. & Vane, J. R. *Nature* 263, 663-665 (1976).

22. Moncada, S. & Vane, J. R. *Br. med. Bull.* 34, 129-135 (1978).
23. Smith, D. R. *et al. Prostaglandins* 18, 423-438 (1979).
24. Dyerberg, J., Jørgensen, K. A. & Arnfred, T. *Prostaglandins* 22, 857-862 (1981).
25. Samuelsson, B. *Biochim. biophys. Acta* 84, 707-713 (1964).
26. Ferretti, A., Schoene, N. W. & Flanagan, V. P. *Lipids* 16, 800-804 (1981).
27. Culp, B. R., Titus, B. G. & Lands, W. E. M. *Prostaglandins Med.* 3, 269-278 (1979).
28. Hornstra, G., Christ-Hazelhof, E., Haddeman, E., ten Hoor, F. & Nugteren, D. H. *Prostaglandins* 21, 727-738 (1981).
29. Falardeau, P., Oates, J. A. & Brash, A. R. *Analyt. Biochem.* 115, 359-367 (1981).
30. Rosenkranz, B. *et al. Biochim. biophys. Acta* 619, 207-213 (1980).
31. Spector, A. A., Kaduce, T. L., Hoak, J. C. & Fry, G. L. *J. clin. Invest.* 68, 1003-1011 (1981).
32. Dyerberg, J. & Jørgensen, K. A. *Artery* 8, 12-17 (1980).
33. Spector, A. A. *et al. J. clin. Invest.* 65, 1003-1012 (1980).
34. Fischer, S., Scherer, B. & Weber, P. C. *Biochim. biophys. Acta* 750, 127-133 (1983).
35. Hamberg, M. & Samuelsson, B. *J. biol. Chem.* 246, 6713-6721 (1971).
36. Fischer, S. & Weber, P. C. *Biochem. biophys. Res. Commun.* 116, 1091-1099 (1983).

Reconstitution with syngeneic plus allogeneic or xenogeneic bone marrow leads to specific acceptance of allografts or xenografts

Suzanne T. Ildstad & David H. Sachs

Transplantation Biology Section, Immunology Branch, Cancer Institute, National Institutes of Health, Building 10, Room 4B17 Bethesda, Maryland 20205, USA

Clinical organ transplantation between genetically disparate individuals currently requires the use of chemotherapeutic agents to suppress the rejection reaction. The deleterious side effects of these reagents and their inability to prevent rejection completely has led to a continuing search for methods to induce specific transplantation tolerance in adult recipients. Numerous experimental animal models utilizing irradiation and bone marrow transplantation coincident with organ transplantation have been proposed¹⁻⁴. Bone marrow transplantation, however, has its own major complications, including graft-versus-host reactions and immunoincompetence, probably resulting from a failure of appropriate immune cell interactions in the reconstituted host^{5,6}. We have now attempted to overcome these difficulties by reconstituting the irradiated host with T-cell depleted bone marrow containing both host (syngeneic) and donor (allogeneic or xenogeneic) components. This technique leads to long-term survival of the reconstituted animals and specific prolongation of subsequent skin grafts of donor type. Animals reconstituted in this fashion are fully reactive to third-party allografts and xenografts and do not appear to manifest signs of graft-versus-host disease.

Inbred mice of the congenic strains C57BL/10Sn (B10) and B10.D2/nSn (B10.D2) (Jackson Laboratory, Bar Harbor) and rats of strains F344 and Wistar Furth (WF) (Charles River Laboratories) were used. Recipient B10 males were given a single dose of 950 rad from a caesium source. Donor bone marrow was treated with rabbit anti-mouse brain antiserum (RAMB) and guinea pig complement to remove mature T lymphocytes as previously described⁷.

Titration were performed to determine the optimal number of syngeneic and allogeneic or xenogeneic bone marrow cells needed to achieve consistent reconstitution and tolerance. The experiments described here were performed with a dose of 5×10^6 syngeneic cells and either 1.5×10^7 allogeneic B10.D2 bone marrow cells or 4×10^7 xenogeneic F344 rat bone marrow cells. Bone marrow preparations were treated separately *in vitro* with RAMB plus complement and were then washed, counted and mixed just prior to inoculation through the tail vein. Animals were subsequently housed in a specific pathogen-free facility at the National Institutes of Health.

The survival of B10 animals reconstituted with B10 plus B10.D2 bone marrow was 100% at 100 days and 88% at 200 days, which is essentially identical to the survival of B10 animals reconstituted with syngeneic bone marrow (data not shown). All irradiation control animals died within 2 weeks. Cytotoxicity

Table 1 Characterization of mice reconstituted with mixed allogeneic bone marrow (B10+B10.D2→B10)

Animal no.	Complement control	% Cytotoxicity		% B10.D2 cells
		B10.D2αB10 + complement*	B10αB10.D2 + complement	
1	3	9	75	92
2	4	80	27	23
3	7	26	80	79
4	7	18	79	87
5	8	32	75	74
6	8	80	25	19
7	4	19	75	82
8	8	70	38	33
9	5	39	38	49
10	8	45	75	64
11	8	57	44	42
12	1	22	52	70
13	5	40	37	48
14	3	55	27	32
15	6	45	43	49
16	4	27	45	64
17	5	92	10	5
Normal B10	4	99	4	0
Normal B10.D2	6	5	100	100

Complement-mediated cytotoxicity was determined on peripheral blood lymphocytes (PBL) by trypan blue exclusion.

* Maximum % lysis by antiserum plus complement. Totals for B10 and B10.D2 do not necessarily equal 100% because of inherent sampling error in the assay. % B10.D2 was calculated by normalizing total specific % kill by the two antisera to 100% after correction for % lysis by complement alone

$$(\% \text{ B10.D2} - C / (\% \text{ B10.D2} + \% \text{ B10}) - 2(C)) \times 100$$

typing of peripheral blood lymphocytes⁸ of reconstituted animals demonstrated persistence of both host- and donor-type lymphoid cells at 60–120 days following reconstitution (Table 1). These animals therefore represent mixed allogeneic chimaeras.

Animals were grafted with full thickness skin grafts from donor and third-party allogeneic strains of mice approximately 12 weeks after reconstitution, using a modification of the method of Billingham⁹. Rejection was considered complete when no clearly viable graft could be detected. Eighty-six per cent of mixed allogeneic chimaeras were tolerant to donor-type skin grafts, showing complete survival of fully viable skin grafts for more than 145 days (Fig. 1). The tolerance was specific, since third-party allogeneic skin grafts were rejected with a median survival time essentially the same as that of unirradiated controls.

B10 recipient mice reconstituted with a mixture of syngeneic and xenogeneic F344 rat bone marrow cells also demonstrated long-term survival similar to that of syngeneically reconstituted animals (86% survival at 100 days). In contrast to the mixed allogeneic chimaeras, peripheral blood lymphocyte typing of mixed xenogeneically reconstituted animals at day 60–70 follow-

Table 2 Characterization of mice reconstituted with mixed xenogeneic bone marrow (B10+rat→B10)

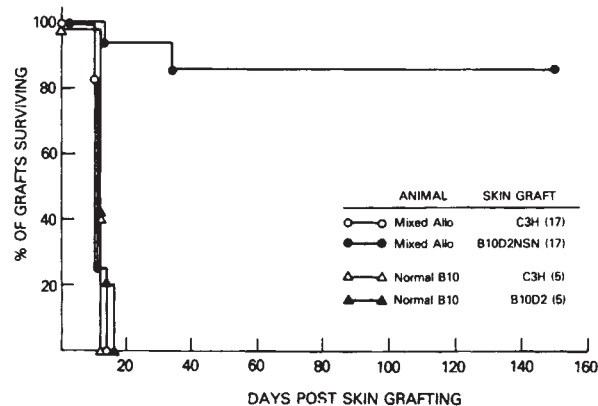
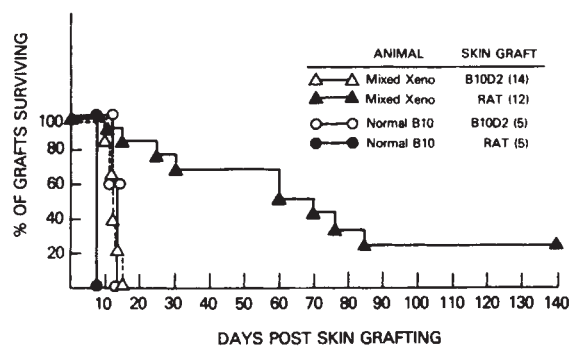
Animal no	% Rat lymphocytes in PBL assessed by complement-mediated cytotoxicity assay*	% Rat lymphocytes in spleen assessed by FACS analysis lymphoid cells†
226	-2	+0.47
227	0	+0.92
228	+1	+0.11
229	+9	+0.87
230	0	-0.29
234	0	+0.94
235	+6	+2.26
250	+5	+2.12
214	0	NT
217	0	NT
240	+2	NT
241	-1	NT
Normal B10 control	0	+0.44±0.03

* % Lysis was that obtained by mouse anti-rat serum plus complement minus complement control.

† % Of cells with fluorescence intensity greater than 300 on log scale following staining with mouse anti-rat serum and fluorescein isothiocyanate conjugated goat F(ab)₂ anti-mouse IgG preparation (FITC-GAMIG) minus % staining with FITC-GAMIG alone. Animals were subjected to partial splenectomy, and splenic lymphoid suspensions were stained with mouse anti-rat antiserum and a FITC-GAMIG. NT, not tested.

ing reconstitution revealed low numbers, if any, of donor lymphoid cells in the peripheral blood of these animals (Table 2). Further analysis of these animals was performed by flow microfluorimetry (FMF) using the fluorescence-activated cell sorter (FACS)¹⁰. Control experiments showed that 0.5% contaminating rat lymphoid cells could reliably be detected. The number of rat lymphoid cells detectable in spleens of mixed xenogeneically reconstituted mice was very small, ranging from 0–2.26% (Table 2).

Nevertheless, these animals appeared to be specifically hyporeactive to donor-type F344 rat skin grafts (Fig. 2). Fourteen B10 animals reconstituted with mixed syngeneic and xenogeneic F344 rat bone marrow cells were grafted with both F344 and B10.D2 skin grafts 75–85 days following irradiation and reconstitution. Rat skin grafts exhibited signs of inflammation beginning 2–10 weeks following grafting. This inflammation often subsided, but eventually recurred in all animals and led either to late rejection or contraction of the skin graft to a small but viable residual. However, all of the animals showed prolonged survival of the rat skin grafts (median survival time [MST] > 50 days) compared with normal B10 animals, which

**Fig. 1** Survival of B10.D2 and C3H full thickness tail skin grafts on lateral thoracic wall of mixed allogeneic chimaera (B10+B10.D2→B10) and controls calculated by life table method.**Fig. 2** Survival of F344 rat and B10.D2 full thickness tail skin grafts on lateral thoracic wall of mixed xenogeneic chimaeras (B10+F344 rat→B10) and controls calculated by life table method. Fourteen reconstituted animals were grafted with both rat and B10.D2 skin grafts, but only 12 of the rat skin grafts were technically successful and could be evaluated.

rejected F344 rat skin grafts in less than 9 days. Furthermore, these same animals rejected the third party allogeneic B10.D2 skin grafts with an MST identical to that of normal B10 controls. This result has subsequently been confirmed in two additional groups of mice. In one of these groups, an additional skin graft was performed from a WF rat, a rat strain differing from the F344 at both major and minor histocompatibility loci¹¹. The WF skin grafts were all rejected promptly, while the F344 skin grafts again showed the same prolonged survival.

Numerous previous studies^{7,12,13} have shown that elimination of mature T cells from donor bone marrow inocula can lead to long-term survival of fully allogeneic bone marrow chimaeras. However, the survival of such animals is generally not as good as that of syngeneically reconstituted controls, and the health of these animals is compromised, probably as the result of partial immunoincompetence^{5,6}. The model for fully xenogeneic chimaeras¹⁴ has been less extensively studied than that of allogeneic chimaeras. In our experience, such animals show prolonged survival over irradiation controls, but all succumb within a few weeks following reconstitution. It is not clear whether the death of these animals is due to a chronic graft-versus-host disease, or to infectious complications secondary to partial immunoincompetence.

In contrast, the studies presented here indicate that reconstitution with mixtures of syngeneic bone marrow and either allogeneic or xenogeneic marrow leads to long-term survival of recipients equivalent to that of syngeneically reconstituted animals. The animals have remained healthy and have shown no evidence of infections or of graft-versus-host disease for longer than 3 months.

Typing of mixed allogeneically reconstituted animals at 2–3 months following reconstitution showed them to be truly mixed chimaeras, bearing variable percentages of syngeneic and allogeneic lymphoid cells in their peripheral circulation. Such animals are immunocompetent¹³ and specifically tolerant to donor skin grafts. In mice reconstituted with mixed syngeneic and xenogeneic bone marrow, we have found little evidence for persistence of the xenogeneic lymphoid compartment. However, these animals also appear to be specifically hyporeactive to subsequent skin xenografts. They may be more similar to neonatally induced allochimaeras¹⁵, in which only very small numbers of persistent donor-type cells can be found¹⁶. Since the survival of skin grafts is frequently more difficult to prolong than that of most other foreign tissues¹⁷, it seems likely that animals manifesting this level of hyporeactivity to skin grafts would be highly tolerant to xenografts of kidneys, hearts and other solid organs.

These data therefore indicate that reconstitution of lethally irradiated mice with mixtures of syngeneic and either allogeneic or xenogeneic bone marrow leads to the induction of specific immune unresponsiveness to subsequent donor skin grafts. The mechanism of this specific unresponsiveness may not be the same in these mixed allogeneic and xenogeneic models, and further mechanistic studies both *in vitro* and *in vivo* are presently in progress. Both types of reconstitution offer models which may be directly applicable to organ grafting in other species, including man.

We thank Dr Hugh Auchincloss Jr for helpful discussions, Mr David Stephany for performing the FACS analysis, Ms Judith Jaworek for manuscript preparation, and Ms Susan O. Sharrow and Drs David Tollerud, Richard Hodes and Mark Peskovitz for review of the manuscript.

Received 30 June; accepted 19 October 1983.

- Rapaport, F. T. *et al. Transplant Proc.* **9**, 891–894 (1977).
- Slavin, S., Strober, S., Fuks, Z. & Kaplan, S. *J. exp. Med.* **146**, 34–48 (1977).
- Dittmer, J. & Bennett, M. *Molec. cell. Biochem.* **21**, 83–94 (1978).
- Myburgh, J. A., Smit, J. A., Browde, S. & Hill, R. H. *Transplantation* **29**, 401–404 (1980).
- Zinkernagel, R. M., Althage, A., Callahan, G. & Welsh, R. M. *J. Immun.* **124**, 2356–2365 (1980).
- Rayfield, L. S. & Brent, L. *Transplantation* **36**, 183–189 (1983).
- Auchincloss, H. Jr & Sachs, D. H. *Transplantation* **36**, 436–441 (1983).
- Arn, J. S., Riordan, S. E., Pearson, D. & Sachs, D. H. *J. Immun. Meth.* **55**, 141–153 (1982).
- Billingham, R. E. in *Transplantation of Tissues and Cells* (eds Billingham, R. E. & Silvers, W. K.) (Wistar Institute, Philadelphia, 1961).

- Sharrow, S. O., Mathieson, B. J. & Singer, A. *J. Immun.* **126**, 1327–1335 (1981).
- Gasser, D. L. *Adv. Immun.* **25**, 93–139 (1977).
- Krown, S. E., Coico, R., Scheid, M. P., Fernandez, G. & Good, R. A. *Clin. Immun. Immunopath.* **19**, 268–283 (1981).
- Singer, A., Hathcock, K. S. & Hodes, R. J. *J. exp. Med.* **153**, 1286–1301 (1981).
- Muller-Ruchholtz, W., Muller-Hermelink, H. K. & Wottge, H. U. *Transplant Proc.* **11**, 517–521 (1979).
- Billingham, R. E., Brent, L. & Medawar, P. B. *Nature* **172**, 603 (1953).
- Streilein, J. W. & Klein, J. *J. Immun.* **119**, 2147–2150 (1977).
- Boyse, E. A., Lance, E. M., Carswell, E. A., Cooper, S. & Old, L. J. *Nature* **227**, 901–903 (1970).

Genetic evidence of X–Y interchange in a human XX male

Albert de la Chapelle*, Patricia A. Tippet†, Gunilla Wetterstrand‡ & David Page§

* Department of Medical Genetics, University of Helsinki, 00290 Helsinki 29, Finland

† MRC Blood Group Unit, Wolfson House, University College London, 4 Stephenson Way, London NW1 2HE, UK

‡ Jorv Hospital, 02740 Esbo 74, Finland

§ Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Of the hypotheses put forward to explain why occasional individuals with two X chromosomes are nonetheless male, the one that has attracted most attention is the possibility¹ that one of the X chromosomes has obtained a small piece of Y chromosome which is sufficient to produce 'maleness'. This hypothesis was based primarily on the observation that in two families with XX males^{2–4} both fathers were Xg(a+) and both probands Xg(a–). (Xg shows X-linked dominant inheritance.) This theory holds that an anomalous X–Y interchange at meiosis in the father resulted in the paternal X chromosome's losing the Xg gene and acquiring a male-determining gene from the Y chromosome. While, for example, the frequencies of Xg phenotypes among XX males^{5,6} and the cytogenetic observation of a structural abnormality in one X^{7,8} are compatible with this hypothesis, direct evidence of it is lacking. Here we describe an XX male who expresses his father's allele for 12E7, a Y-linked marker, but fails to express his father's allele for Xg, an X-linked marker. These findings strongly suggest that anomalous X–Y interchange occurred in this case and perhaps in that of many other XX males. We suggest that a male-determining gene on the Y has also been translocated to the X and caused maleness in the proband. These results are discussed in the light of current models of X–Y chromosomal homology.

A boy born in 1980 had slightly dysmorphic facial features but was otherwise normal. He has remained in good health but his physical growth is greatly retarded (at 2 yr and 3 months old, height = 75 cm, weight = 7.4 kg). His height development is well below the 2.5 percentile for boys and even for girls. Clinical and laboratory investigations had failed to disclose anything abnormal to account for the retarded growth until a karyotype showed him to be 46, XX; there was no evidence of mosaicism in the two tissues karyotyped (lymphocytes and skin fibroblasts). The parents (born in 1953) are non-consanguineous. All four grandparents are alive. The proband has a brother born in 1978 and a sister born in 1983. All family members are healthy.

All family members (Fig. 1) were tested for the following blood groups: ABO, MNSs, P₁, Rh, Lu^a, Kell, Lewis, Duffy, Kidd, Dombrock and Co^b. That II-1 is indeed the father of III-2 is supported by the inheritance of these markers (data not shown). Figure 1 shows the segregation of the blood group Xg, the antigen 12E7, and the polymorphic X–Y homologous DNA locus designated DXYS1 (ref. 9) in the members tested. Both parents, II-1 and II-3, are Xg(a+); the mother is heterozygous Xg^aXg as her father, I-3, is Xg(a–). Both the proband, III-2, and his brother, III-1, are Xg(a–) so they have received the maternal Xg gene. All Xg(a+) individuals express high levels

of 12E7 (ref. 10). However, in families with Xg(a-) members, the quantitative polymorphism of 12E7 expression shows Y-linked inheritance¹¹. For example, 77 Xg(a-) high-level 12E7 males (first tested member of sibship) have 110 Xg(a-) younger (or tested later) brothers of whom 107 have high-level 12E7 and 3 have low-level 12E7 expression; 18 Xg(a-) low-level 12E7 males have 19 Xg(a-) younger brothers of whom 17 have low-level 12E7 and 2 have high-level 12E7 expression. These comparisons give $\chi^2 = 20.51$ and 15.81 respectively if 12E7 were not Y-controlled. The five exceptional brothers could be the result of X-Y interchange, or of difficulties in testing travelled red cells, or of a different father (the father of only one of the exceptions has been tested)¹¹.

In the family reported here, both the paternal uncle, II-2, and elder brother, III-1, are Xg(a-) and high expressors of 12E7, suggesting that the Y chromosome in the paternal family carries Yg^a. (The Y-linked 12E7 locus has been designated Yg on the basis of its hypothesized homology to Xg¹⁰. The allele for high expression of 12E7 has been designated Yg^a.) The maternal grandfather, I-3, is Xg(a-) low 12E7 and, therefore, XgYg.

The proband does not carry the paternal Xg^a allele but has received a gene producing high 12E7 expression. As this could not have come from the mother, the most straightforward explanation is an interchange of the Xg and Yg loci in paternal meiosis. As a consequence of this, the X chromosome of paternal origin has lost its Xg^a but acquired Yg^a. The presence of testes and male differentiation in the proband is probably the consequence of a sex-determining gene having been exchanged as well. Perhaps a male-determining gene normally located in proximity to Yg on the Y chromosome has been translocated to the X. Alternatively, an X-borne gene involved in sex determination may have been lost or inactivated.

Our interpretation assumes that one of the proband's X chromosomes is indeed paternal. The TaqI restriction fragment length polymorphism at the DXYS1 locus shows normal segregation of the Y-specific 15-kilobase (kb) fragment with the Y chromosome in the father, II-1, and brother, III-1 (Fig. 1). The proband has a double dose of the X-specific 11-kb allele (one from the father and one from the mother). This is not formal proof of the paternal derivation of one X, as it is also consistent with the proband having two copies of the maternal X carrying the 11-kb allele (and Xg). However, the latter possibility is inconsistent with the 12E7 data. The segregation of eight additional X-linked DNA polymorphisms was examined in this family but none of the polymorphisms proved to be informative with respect to a paternal contribution (data not shown).

It has been shown recently that many XX males have one paternal and one maternal X chromosome. When the segregation of the DXYS1 alleles was tested in six families, four were informative. In each of these four cases the XX male proband was heterozygous^{11,12} while the mother was homozygous for one of the alleles and the father hemizygous for the other (D.P. and A. de la C., unpublished results). This supports the notion that one X is paternal in many, if not most, XX males.

The proposed interchange explains three findings (loss of Xg^a, acquisition of Yg^a, maleness) as the consequence of one event. The alternative—non-disjunction at maternal meiosis plus non-disjunction at the first paternal meiotic division or zygotic loss of the Y—does explain the chromosomal and Xg data but fails to explain the level of 12E7 expression and the maleness of the proband. Hence, the interchange mechanism must be considered the most likely one.

At present we do not know what proportion of XX males arises by the type of interchange that occurred in this case. That it may be common is suggested by the fact that at least nine families are known in which the XX male proband apparently did not inherit his father's Xg allele⁶. Furthermore, the frequency of Xg phenotypes in XX males is close to that of XY males and is significantly different from that of females (R. Sanger and P.A.T., unpublished observations on 94 northern European XX males).

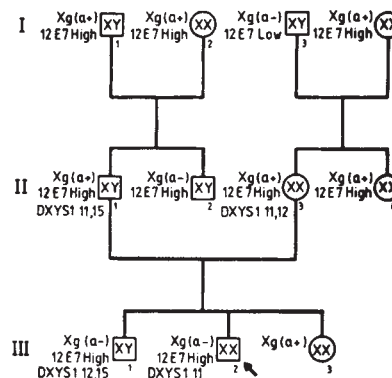


Fig. 1 Pedigree of the family showing the sex chromosome constitution and segregation of the blood group Xg, the antigen 12E7, and the X-Y homologous DNA locus DXYS1.

In contrast, at least five families are known in which the XX male proband did inherit his father's Xg allele⁶. These cases either arose by another mechanism or the interchange involved different regions of those chromosomes. Light and electron microscopic evidence^{12,13} indicates that terminal portions of the short arms of the human X and Y chromosomes normally pair at male meiosis. It has been suggested that crossing-over within this 'pairing region' occurs as a normal event^{1,14-16}, but there is no genetic evidence for this. Here we have described a clearly anomalous meiotic interchange of normally X-linked and Y-linked markers. Of note is the fact that the Xg gene maps to the 'pairing region' on the X¹⁷ while a 12E7 structural gene maps to the euchromatic portion of the Y¹⁸, which includes the pairing region. Thus, the X-Y interchange reported here may be an example of a common event, but one which is abnormally proximal to the centromere.

The aetiology of XX males is probably heterogeneous. The underlying mechanisms may include errors of X-chromosome inactivation, mutations in autosomal or X-chromosomal genes, and X-Y interchange (probably with different breakpoints in different cases). A somewhat analogous finding is that the sex reversal mutation in the mouse (*sxr*) is now known to be due to a complicated Y-X exchange^{19,20}.

This study was aided by grants from the Sigrid Jusélius Foundation, the Nordisk Insulinfond, the Academy of Finland, the NIH (USA), and the Wills Foundation. We thank Dr Peter Goodfellow for the anti-12E7 antibody and Drs David Botstein and Arlene Wyman for critical reading of the manuscript.

Received 1 August; accepted 1 November 1983.

1. Ferguson-Smith, M. A. *Lancet* **ii**, 475-476 (1966).
2. de la Chapelle, A., Hortling, H., Niemi, M. & Wennström, J. *Acta med. scand. Suppl.* **412**, 25-38 (1964).
3. de la Chapelle, A., Hortling, H., Wennström, J., Niemi, M. & Johansson, C.-J. *Acta endocr. Suppl.* **100**, 90 (1965).
4. de la Chapelle, A., Similä, S., Lanning, M., Konturi, M. & Johansson, C.-J. *Hum. Genet.* **11**, 286-294 (1971).
5. Race, R. & Sanger, R. *Blood Groups in Man* 6th edn (Blackwell, Oxford, 1975).
6. de la Chapelle, A. *Hum. Genet.* **58**, 105-116 (1981).
7. Evans, H. J., Buckton, K. E., Spowart, G. & Carothers, A. D. *Hum. Genet.* **49**, 11-31 (1979).
8. Magenis, R. R. *et al. Hum. Genet.* **62**, 271-276 (1982).
9. Page, D. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 5352-5356 (1982).
10. Goodfellow, P. N. & Tippet, P. *Nature* **289**, 404-405 (1981).
11. Tippet, P., Shaw, M.-A., Daniels, G. L. & Green, C. A. *Human Gene Mapping Vol. 7* (in press).
12. Pearson, P. C. & Bobrow, M. *Nature* **226**, 959-961 (1970).
13. Moses, M. J., Counce, S. J. & Paulson, D. F. *Science* **187**, 363-365 (1975).
14. Polani, P. E. in *Mechanisms of Sex Differentiation in Animals and Man* (eds Austin, C. R. & Edwards, R. G.) 465-488 (Academic, London, 1981).
15. Polani, P. E. *Hum. Genet.* **60**, 207-211 (1982).
16. Burgoyne, P. S. *Hum. Genet.* **61**, 85-90 (1982).
17. Ferguson-Smith, M. A., Sanger, R., Tippet, P., Aitken, D. A. & Boyd, E. *Cytogenet. Cell Genet.* **32**, 273-274 (1982).
18. Goodfellow, P. *et al. Nature* **302**, 346-349 (1983).
19. Singh, L. & Jones, K. W. *Cell* **28**, 205-216 (1982).
20. Evans, E. P., Burtenshaw, M. D. & Cattanach, B. M. *Nature* **300**, 443-445 (1982).

Human XX males with Y single-copy DNA fragments

Georges Guellaen*, Myriam Casanova†, Colin Bishop*†, Danielle Geldwerth*, Gabriel Andre‡, Marc Fellous† & Jean Weissenbach*

* Unité de Recombinaison et Expression Génétique, INSERM U 163, CNRS LA 271 and † Unité d'Immunogénétique Humaine, Institut Pasteur, 28 rue du Dr Roux, 75015 Paris, France
‡ Hospices Civils, 67000 Strasbourg, France

In humans, XX maleness is the best known example of a sex reversal syndrome occurring with an incidence of one XX male among approximately 20,000 to 30,000 newborn boys¹⁻⁴. The karyotypes of the majority of these individuals are apparently normal, with respect to the numbers and structure of the chromosomes, but is in contradiction with the phenotypic sex which they display. XX maleness may be either a non Y-related mechanism triggered by a mutation on another chromosome³⁻⁵ or could be the result of the expression of some cytogenetically undetectable Y chromosome material present in the genome of such individuals. Recently, a number of human Y-specific single copy probes have been isolated^{6,7}. In this study, using several of these Y-specific probes we definitively demonstrate the presence of Y-chromosomal material in the genome of some 46,XX human males. These XX males carry only a fraction of the human Y chromosome. In the three positive cases reported here, presence of inclusive overlapping chromosomal fragments has been detected, implying a genetic heterogeneity of these patients.

The 4 cases of XX male analysed are sporadic with no familial incidence of genetic intersexuality. Their clinical and biological features are summarized in Table 1. DNAs from cloned lymphoblastoid cell lines or lymphocytes originating from the four different human XX males have been digested with *Eco*RI, *Hind*III or *Taq*I, fractionated and blotted onto DBM-cellulose paper⁸. The DBM papers were successively hybridized to different probes detecting Y-specific DNA fragments. Under stringent hybridization conditions these probes detected single copy DNA fragments whereas under less stringent conditions additional bands are sometimes observed that are not strictly homologous. In order to detect a maximal number of Y-specific bands, low stringency washing conditions have been used in this survey whenever possible. Figure 1 shows the hybridization pattern of two probes 52d and 50f2 to *Taq*I digests. Probe 52d (Fig. 1A) detects three Y-specific bands in normal XY DNAs and a small X-located fragment⁷ not present in this figure. Two out of the three Y-specific fragments are also present in XX males 1 and 4 indicating the presence of Y chromosomal DNA in some XX males. Since the third Y-specific band does not appear in these cases (1 and 4), we conclude that the Y chromosome is not present as a whole. Confirmation of this result is obtained with probe 50f2 (Fig. 1B). This probe detects 1 autosomal band at 12 kilobases (kb), in both male and female DNA, and in addition, four Y-specific fragments. Here again, only 2 out of 4 of the Y specific bands are present in the same patients 1 and 4.

The data on the Y-specific sequences, tested as probes on genomic blots of XX males, are summarized in Table 2. Out of the four individuals, only three reacted with the probes used (case 1, 3 and 4). The three positive cases did not react to the same extent with the probes tested. Although case 1 and 4 reacted with six and five probes respectively, case 3 reacted only with probe 47c. This observation implies that these XX male patients each contain a different amount of Y-chromosomal DNA and thus exhibit a genetic heterogeneity. In all three cases, the same subset of Y-specific bands was detected, implying presence of overlapping chromosomal fragments. The genome

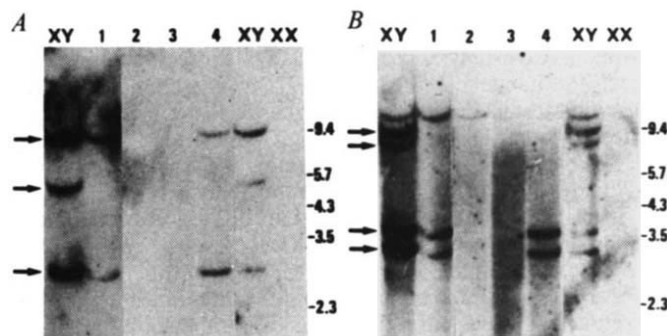


Fig. 1 DNA was extracted as previously described^{6,7} from lymphocytes of normal individuals or from cloned lymphoblastoid cell lines derived by Epstein-Barr virus transformation of lymphocytes from XX males. The probes were derived from a partial Y chromosome library as previously reported^{6,7}. The genomic blot used in A and B was prepared by digesting genomic DNA to completion with *Taq* (Bio Labs) used according to manufacturers' specifications, separating on 1% agarose gels and transferring to DBM paper. Normally 15 µg of DNA was used per lane. The DBM paper blots were hybridized⁸ with 30 × 10⁶ c.p.m. per blot of the probes 52d (A) or 50f2 (B) (10⁸ c.p.m. µg⁻¹) and washed under non-stringent conditions (2 × SSC; 68 °C; 2 × 60 min). Exposure of the film was done for various times at -70 °C in the presence of an intensifying screen. XY: normal male; XX: normal female; 1-4: different XX males. The male specific bands are arrowed on the left of the two panels. The size of the markers is indicated in kilobase pair units on the right.

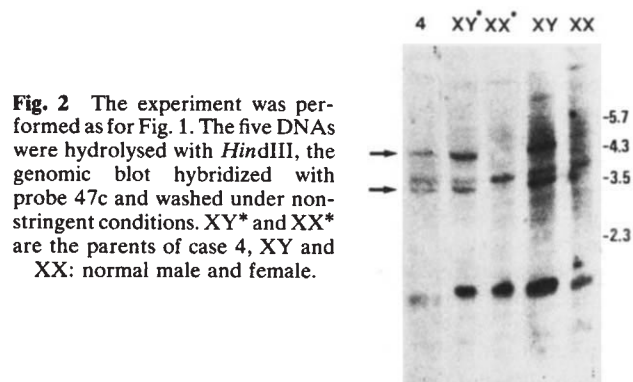


Fig. 2 The experiment was performed as for Fig. 1. The five DNAs were hydrolysed with *Hind*III, the genomic blot hybridized with probe 47c and washed under non-stringent conditions. XY* and XX* are the parents of case 4, XY and XX: normal male and female.

of case 1 contains the most numerous Y DNA fragments. These fragments completely overlap with the fragments present in case 4 which themselves overlap with the fragments of case 3. This totally inclusive overlapping could be observed by chance, but it is also consistent with an involvement of one of the telomeric parts of the Y chromosome. If, in addition, one assumes that only a single Y-chromosomal piece is present in the genome of these XX males, blocks of overlaps can be positioned with reference to a common origin and the distance of a DNA fragment from this origin is inversely related to the frequency of this fragment in XX males. This would represent the first attempt to order Y DNA fragments by deletion mapping.

Since the probes described here also hybridize to non-Y material, the possibility exists that the fragments detected in XX males are due to a restriction fragment length polymorphism of this non-Y material. This possibility can be rejected in view of the following arguments: (1) the non-Y bands are still present in the expected amount in the XX males according to hybridization intensity, (2) all the putative Y bands in the XX males appear exactly at the same level as in normal males (Fig. 1) and in different restriction digests (not shown). It is practically impossible that, by chance, restriction length polymorphism could have resulted in numerous bands all of the exact same size as the Y-specific ones.

The parental origin of the Y-specific fragment could be assessed further for case 4. Figure 2 represents a genomic blot

Table 1 Clinical and biological summary of the 4XX males analysed

	Case 1	Case 2	Case 3	Case 4
Age (yr)	21	11	22	20
Sterility	Azoospermy	NT	Azoospermy	Azoospermy
Testes	Small	Extremely reduced	Small	Small
Gynaecomasty	Absent	Absent	Absent	Absent
Histopathology of the gonad	NT	Testicular structure without germ cells, with Sertoli cells	Testicular structure without germ cells, with Sertoli cells	Testicular structure without germ cells, with Sertoli cells
FSH level	High	High	High	High
Testosterone level	Low	Very low	Low	Low
Karyotype	XX	XX	XX	XX
H-Y	Present	Present	Present	Present

Karyotype and H-Y determination¹⁵ were performed on lymphoblastoid cell lines and peripheral lymphocytes. Similar results were obtained with both types of cells. NT, not tested. In all the four cases, at least 200 mitoses were analysed both on peripheral lymphocytes and lymphoblastoid cell line obtained by Epstein-Barr virus transformation of B lymphocytes. Trypsin band (G banding) and Q banding were performed. No Y material could be detected and both X chromosomes appear of normal size. Familial blood groups, including HLA antigen were performed. No exclusion of paternity was observed. Xg anomalous segregation were found for cases 2 and 3 and will be described elsewhere. For the two other cases, the Xg phenotype was not informative.

of *Hind*III restriction enzyme hydrolysis of case 4 along with DNA from his parents (asterisk) and normal male and female DNA. Hybridization to probe 47c reveals an identical pattern between case 4, XY* and XY, and a total absence of male specific bands on the DNA of the mother(XX*). This result excludes maternal inheritance of the Y-chromosomal DNA in this case.

It is generally thought that the mammalian Y chromosome is the bearer of a male sex determining locus^{1,2}. It is reasonable to conclude that this putative locus is present in the piece of Y chromosome detected in XX male patients. In case 3 only a single Y specific fragment has been detected. This suggests that only a small part of the Y chromosome is present and that this part is sufficient for controlling essential steps of the male gonadal sex determination.

This study represents the first unambiguous molecular evidence for the presence of Y DNA material in some human XX male genomes. Furthermore, it also proves that only a part of this chromosome is present and this is probably the reason for the absence of its detection by cytogenetical means. These results can be related to the recent explanation of the sex reversal mutation in mouse^{9,10}. However, it should be pointed out that in man this syndrome is mainly sporadic, whereas in mouse, the *Sxr* mutation is genetically transmissible by affected XY males¹¹. In the aetiology of human XX maleness, the Y chromosome is mainly involved in two hypotheses: translocation of a piece of Y chromosome and hidden mosaicism. Since the number of mitoses usually examined in karyotyping studies is relatively important, the cytogenetical approaches would certainly remain more sensitive for the detection of a mosaicism.

The form in which this Y chromosome DNA is present in XX males is at present unknown. The most probable explanation involves a translocation event and as reported elsewhere, the X chromosome is the most likely candidate¹²⁻¹⁴. X-Y interchange can even be suspected in case 3 where an anomalous segregation of the Xg blood marker has been observed (M. Fellous, unpublished results).

One patient (case 2) remained negative in this survey, it is not known if this is related to the fact that this case displayed the most severe hypogonadism. Since it is very difficult to know if all the regions of the Y chromosome are covered by our probe collection, presence of Y material cannot therefore be excluded in negative cases for the moment. In the near future, it should be possible, with the availability of more Y-specific probes, to discriminate unambiguously between putative true XX males and XX males bearing Y DNA material.

We thank Professor P. Tiollais for his support, C. Johnsson and A. Abadie for technical assistance, F. Latron for the karyotyping and H-Y studies; DNA samples and DBM blots have

Table 2 Male specific bands in the genome of XX males

Probe	Normal male	XX males		
		1	3	4
47c	2(T)	2	2	2
50f2	4(T)	2	—	2
52d	3(T)	2	—	2
48	1(H)	1	—	1
118	8(T)	4	—	4
27	1(T)	1	—	—
64a7	1(E)	—	—	—
37c	1(E)	—	—	—
49f	3(T)	—	—	—
12f3	1(T)	—	—	—
Male repetitive* probe	1(E)	—	—	—

DNA, genomic blots and hybridization procedures are as depicted in the legend of figure 1. The results refer to the number of male specific bands present on a *Taq* (T), a *Hind*III (H) or an *Eco*RI (E) blot. Following hybridization, a low stringency of washing was used (see Fig. 1) except for probes 64a7, 37c, 49f, and the repetitive probe which were hybridized under high stringency (0.1×SSC instead of 2×SSC).

* Repetitive probe refers to male specific 3.4 kb *Hae*III fragment¹⁶.

been shared with J. L. Mandel whose help is gratefully acknowledged; we are also very grateful to Professor Roy and Dr Caporal for their help and collaboration; C.E.B. was in receipt of a Joliot-Curie grant. This work was supported by INSERM grant PRC 135012 and CRL 814030, DGRST 870346 and by the Fondation de la Recherche Médicale.

Received 15 August; accepted 19 October 1983.

- Haseltine, F. P. & Ohno, S. *Science* **211**, 1272-1278 (1981).
- Gordon, J. W. & Ruddle, F. H. *Science* **211**, 1265-1271 (1981).
- Polani, P. E. in *Mechanisms of Sex Differentiation in Animals and Man* (eds Austin, C. R. & Edwards, R. G.) 465-529 (Academic, New York, 1981).
- de la Chappelle, A. *Hum. Genet.* **58**, 105-116 (1981).
- Wolf, U. *Hum. Genet.* **58**, 25-28 (1981).
- Bishop, C. E. *et al. Nature* **303**, 831-832 (1983).
- Bishop, C. E., Guellaen, G., Geldwerth, D., Fellous, M. and Weissenbach, J. *J. mol. Biol.* (in the press).
- Alwine, J. C., Kemp, D. J. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5350-5354 (1977).
- Singh, L. & Jones, K. W. *Cell* **28**, 205-216 (1982).
- Evans, E. P., Burtenshaw, M. D. & Cattanaach, B. M. *Nature* **300**, 443-445 (1982).
- Cattanaach, B. M., Evans, E. P., Burtenshaw, M. D. & Barlow, J. *Nature* **300**, 445-446 (1982).
- Ferguson-Smith, M. A. *Lancet* **11**, 475-476 (1966).
- Evans, H. J., Buckton, K. E., Spowart, G. & Carothers, A. D. *Hum. Genet.* **49**, 11-31 (1979).
- Magenis, R. E. *et al. Hum. Genet.* **62**, 271-276 (1982).
- Casanova-Bettane, M., Latron, F., Jakob, H. & Fellous, M. *Hum. Genet.* **58**, 21-24 (1981).
- Cooke, H. *Nature* **262**, 182-186 (1976).

The establishment and maintenance of pregnancy using *in vitro* fertilization and embryo donation in a patient with primary ovarian failure

Peter Lutjen, Alan Trounson, John Leeton, Jock Findlay*, Carl Wood & Peter Renou

Department of Obstetrics and Gynaecology, Monash University, Queen Victoria Medical Centre and *Medical Research Centre, Prince Henry's Hospital, Melbourne, Australia 3000

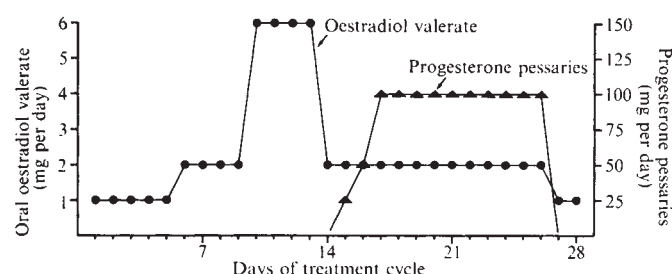


Fig. 1 Dose regime for the cycle replacement of oestrogen and progesterone. Oestradiol valerate was taken orally at 24-h intervals, at doses of 1 and 2 mg per day (days 1-9 and 14-28) or at 8-h intervals at a dose of 6 mg per day (days 10-13). Progesterone was administered as a single 25 or 50 mg intravaginal pessary once a day (days 15-16) or two pessaries introduced separately at 12-h intervals (days 17-26).

Ovarian steroid replacement therapy in the ovariectomized ewe, given in the correct sequence to mimic endogenous steroid changes in the normal ovulatory cycle, allows the development of embryos transferred *in utero*¹⁻³. A similar type of sequential therapy was designed for steroid replacement in women with primary ovarian failure. This produces the histological changes in uterine endometrial morphology and plasma oestradiol and progesterone similar to those observed in the normal ovulatory cycle^{4,5}. We now report that in one of these women a donated oocyte⁶, fertilized by her husband's spermatozoa⁷ and cultured to the two-cell stage *in vitro*, was transferred *in utero*, resulting in a normal pregnancy and the delivery of a healthy child. Oestrogen therapy was withdrawn at 12 weeks and progesterone at 19 weeks gestation. This technique allows the treatment of human infertility due to primary ovarian failure.

The recipient was a 25-year-old married woman complaining of 5 years of amenorrhoea following withdrawal from contraception. Endocrine investigations throughout this period had shown continually depressed urinary oestradiol excretion on 24 separate occasions and continually elevated plasma gonadotropins in the postmenopausal range (follicle stimulating hormone, luteinizing hormone > 30 IU l⁻¹), indicative of primary ovarian failure. Treatment with both human chorionic gonadotropin (β -HCG, 3,000 IU per day for 3 days) and a combination of ethinyloestradiol and progesterone failed to elicit any detectable ovarian response in terms of oestradiol production. The patient presented with characteristic 'hot flushes' and vaginal dryness, indicative of her hypo-oestrogenic state. Secondary sexual development appeared normal and her karyotype (G banded) was normal (46 XX). Both thyroid and adrenal functions were assessed and were normal. Diagnostic laparoscopy at 20 years

of age revealed marked ovarian atrophy and complete absence of follicular development. An ovarian biopsy showed an absence of primordial follicles. At 25 years of age the patient commenced ovarian steroid replacement with the intention of receiving a donated embryo.

The regimen of ovarian steroid replacement with oral oestradiol valerate (Progynova, Schering) and intravaginal progesterone pessaries is given in Fig. 1. The treatment was formulated to mimic as closely as possible the ovarian steroid profile throughout a normal 28-day ovulatory cycle⁵. On the completion of the 28th day of steroid replacement, the patient recommenced therapy at day 1 and repeated this treatment in a cyclic manner.

Peripheral plasma oestradiol and progesterone levels were monitored during the third treatment cycle and on the 16th day a single two-cell embryo was transferred to the patient⁸. Oestrogen and progesterone replacement was continued, as shown in Figs 2 and 3. Following embryo transfer, progesterone pessaries were replaced by intramuscular progesterone injections (Proluton, Schering) for the remaining period of steroid therapy.

Using procedures described previously for embryo donation⁶, a single oocyte was inseminated with spermatozoa prepared from a semen sample from the recipient's husband⁷. The donor of the oocyte was a 29-yr-old woman whose infertility was due to bilateral tubal blockage. Five oocytes were obtained from the donor during an *in vitro* fertilization treatment cycle and

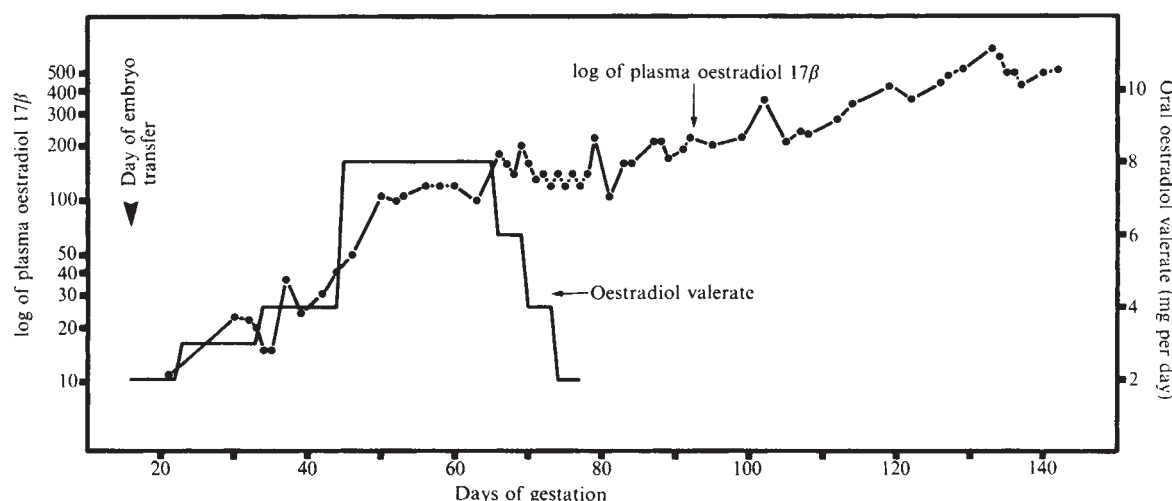


Fig. 2 Oestradiol valerate replacement therapy following embryo transfer and plasma oestradiol-17 β levels (●) during pregnancy. Doses of 2 mg per day were taken at 24-h intervals, 3 and 6 mg at 8-h intervals, and 4 and 8 mg per day taken at 6-h intervals.

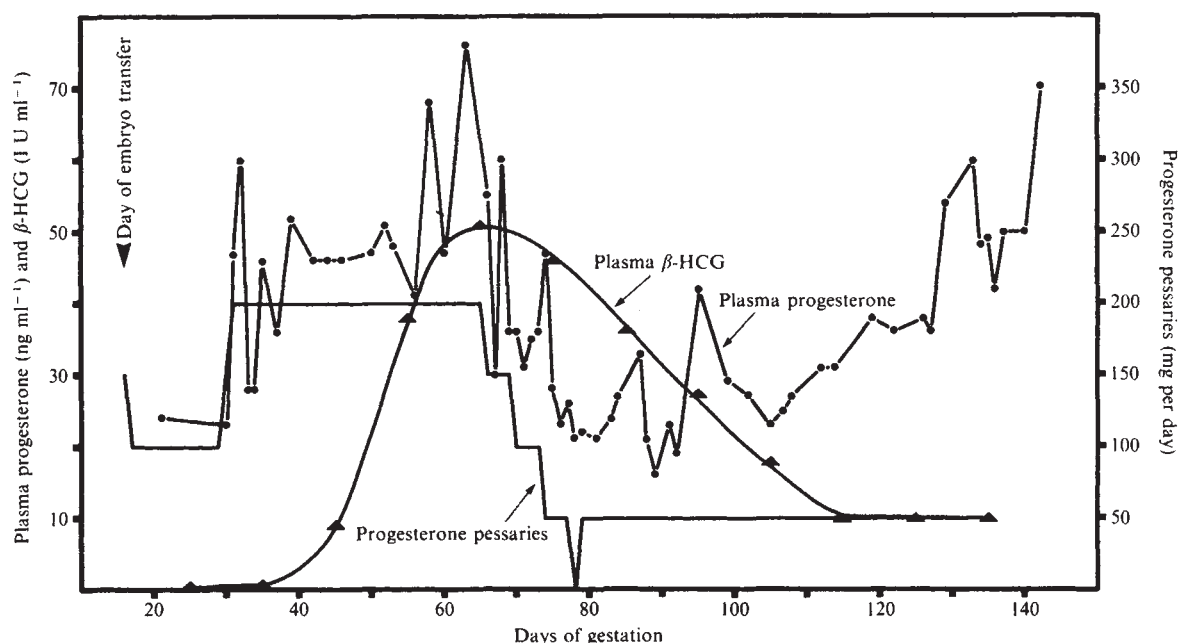


Fig. 3 Intramuscular progesterone (Proluton, Schering) replacement therapy following embryo transfer and plasma progesterone (●) and endogenous β -HCG (▲) levels during pregnancy. Doses of 50 mg per day were given at 24-h intervals, 100 mg per day at 12-h intervals, 150 mg per day of 8-h intervals and 200 mg per day at 6-h intervals.

she and her husband consented to donate one oocyte to the recipient couple. The donor failed to become pregnant on the cycle of IVF treatment after transfer of four of her own two-cell embryos *in utero*.

Pregnancy was confirmed in the recipient by normal changes in β -HCG (Fig. 3). The dosage of oestradiol valerate and progesterone administered was altered to keep the patient's peripheral plasma levels of these hormones in the normal range for pregnancy (Figs 2, 3). Ultrasonic reports on the fetus and placenta at 8, 12, 17 and 28 weeks of gestation were normal. Oestrogen therapy was withdrawn at 12 weeks' gestation. Progesterone therapy was unable to be completely withdrawn until 19 weeks because plasma progesterone levels declined to below the normal levels expected in pregnancy when withdrawal was attempted at 12 weeks.

The pregnancy continued normally until 38 weeks of gestation when a healthy baby boy was delivered by elective caesarean section. HLA-AB typing using peripheral blood lymphocytes in a standard microlymphocytotoxicity assay excluded the recipient as the biological mother and verified the association of the baby with the donor. The establishment and maintenance of pregnancy in the present report confirms our hypothesis⁶ that in women with primary ovarian failure, progesterone and oestradiol can be used successfully as replacement therapy for the establishment and maintenance of pregnancy following transfer of donor embryos. In these women, there is no other treatment for their infertility. Prior to this successful case, we had transferred embryos to six patients, given similar or slightly modified steroid hormone treatments without success.

Shortly after the pregnancy was established, the Premier of the state of Victoria requested the cessation of all work involving donated oocytes to allow consideration of the social, ethical and legal issues of gamete donation in IVF by a state government appointed committee. The committee's recommendations in favour of continuing the work were reported in September 1983⁹. However, the state government has not revoked the moratorium and, as a consequence, we have not been able to continue the research programme and repeat the successful procedure described.

During the preparation of this paper a report was published on the establishment and maintenance of pregnancy in ovariectomized monkeys given ovarian steroid implants¹⁰. Future studies in this area need to determine whether sequential therapy designed to mimic normal endocrine levels in the ovulation cycle is essential for the establishment of pregnancy, as studies in sheep have demonstrated¹⁻³, or whether a simplified steroid therapy regimen may be used. The use of synthetic progestagens instead of progesterone would also simplify hormone replacement therapy for the patients. Pregnancy may be limited to the synchronous transfer of particular cleavage stage embryos on specific days of the treatment cycle. More flexibility of the timing of oocyte donation and embryo transfer would be achieved by embryo cryopreservation¹¹. Legal and ethical difficulties encountered with embryo donation⁶ will be satisfied if governments adopt recommendations such as those proposed in Victoria⁹.

We thank all the members of the Monash-Epworth Infertility Unit and Department of Obstetrics and Gynaecology for helpful advice and assistance, and Dr Brian Tait of the Royal Melbourne Hospital for HLA typing. Progesterone and Proluton were provided by Schering Ltd.

Received 9 December; accepted 16 December 1983.

1. Trounson, A. O. & Moore, N. W. *Aust. J. biol. Sci.* **21**, 511-517 (1974).
2. Miller, B. G. & Moore, N. W. *Aust. J. biol. Sci.* **29**, 563-573 (1976).
3. Moore, N. W., Miller, B. G. & Trappl, M. N. *J. Reprod. Fert.* **68**, 129-135 (1983).
4. Conti, A. *et al. Proc. Aust. Soc. Reprod. Biol.* **14**, 116 (1982).
5. Lutjen, P. J. *et al. Proc. Endocrine Soc. Aust.* **26**, 112 (1983).
6. Trounson, A., Leeton, J., Besanko, M., Wood, C. & Conti, A. *Br. med. J.* **286**, 835-838 (1983).
7. Mahadevan, M. M., Trounson, A. O. & Leeton, J. F. *Fert. Steril.* **40**, 340-343 (1983).
8. Leeton, J., Trounson, A. O., Jessup, D. & Wood, C. *Fert. Steril.* **38**, 156-161 (1982).
9. *Law Reform Commission Report on Donor Gametes in IVF* (Victoria Law Offices, Melbourne, 1983).
10. Hodgen, G. D. *J. Am. med. Ass.* **216**, 2167-2171 (1983).
11. Trounson, A. & Moir, L. *Nature* **305**, 707-709 (1983).

Similar level of polyteny in bands and interbands of *Drosophila* giant chromosomes

Anne Spierer & Pierre Spierer

Department of Molecular Biology, University of Geneva,
30, quai Ernest-Ansermet, 1211 Geneva 4, Switzerland

The giant polytene chromosomes of *Drosophila melanogaster* have long been of interest to the geneticist because of the visible map of the genome provided by their characteristic banding patterns¹. An issue in the understanding of the molecular basis of chromosome banding has been whether the chromosomal DNA is replicated to the same extent in bands and interbands. Although various suggestions have been put forward the point has remained controversial²⁻⁵. We have isolated 315 kilobases (kb) of contiguous *Drosophila* genomic DNA⁶ which spans an interval of approximately 13 bands and interbands of the polytene chromosomes⁷. We report here the measurement of the relative level of DNA replication in polytene chromosomes of 84 adjacent restriction fragments of our cloned DNA. We conclude that there are no significant differences in the level of polyteny within the large band and between bands and interbands of this region. This result supports the 'folded fibre' model of polytene chromosome organization⁸, rather than models involving disproportionate replication along the banding pattern²⁻⁴.

The 315 kb of chromosomal DNA sequence that we have isolated includes 13 ± 1 chromomeric units (band and adjacent interband) of sizes ranging from about 3 to 160 kb, and 12 recessive lethal complementation groups⁷. To detect any local variation in polyteny we decided to measure the relative DNA replication level in polytene chromosomes of 84 fragments representing this interval by quantitative Southern blot hybridizations^{9,10}. Two control experiments checked that this method can quantitate single copy sequences in genomic DNA. A reconstruction experiment demonstrated that the autoradiographic signal is indeed proportional to the relative abundance of a unique restriction fragment in a genomic background (Fig. 1). In a second control, the sensitivity of the method was tested *in vivo*, with a heterozygous deletion ending within the chromosomal walk. Single copy fragments within the deletion produced about half the signal that was seen with wild-type DNA or when compared with probes outside the deletion (Table 1). From

Table 1 Quantitative Southern blot hybridizations distinguish one dose from two doses of a single-copy sequence within genomic DNA

Probe ⁷	Fragment size (kb)	Signal ratio (wt/ry ⁶¹⁴)	Mean of ratios
2189	3.1	0.42	0.46
	1.9	0.39	
	1.5	0.56	
2849	1.7	1.0	0.93
	1.4	0.78	
	1.1	1.0	
Wild-type chromosome	2849		2189
Df(3R)ry ⁶¹⁴ /wt			

2 µg of wild-type genomic DNA and 4 µg of DNA from the heterozygous deletion Df(3R)ry⁶¹⁴ (refs 7 and 13) were digested with *Eco*RI, size-separated on agarose gels and transferred to nitrocellulose. Filters were hybridized to the recombinant 2189⁶ or 2849⁶. The autoradiographic signals were measured by scanning densitometry. The position of the two probes relative to the deletion breakpoint is shown schematically under the table.

these experiments and from tests of reproducibility (not shown), we felt confident that any local change in polyteny larger than twofold from the mean could be detected.

DNA was prepared from polytene tissue (salivary glands from third instar larvae) and from diploid tissues (embryos). Parallel tracks of equal amounts of 'polytene' and 'diploid' DNA digested with restriction enzymes were transferred to nitrocellulose paper and hybridized to a set of probes representing the walk. The signal could thus be expressed for each fragment as a 'polytene/diploid' ratio. A second calibration step was to hybridize all the filters to the same probe, to correct them for unequal loading or differences in transfer efficiency. This probe (2148)⁶ was arbitrarily chosen as providing a polytene/diploid ratio of 1.0. The filters were then recycled and re-used with other probes to span the whole chromosomal walk. Figure 2 shows autoradiographs of eight filters hybridized to eight different segments of the walk and to the calibration segment. We also took the precaution not to quantitate fragments larger than 5 kb, for they are frequently under-represented in DNA preparations of salivary glands because of shearing (our unpublished observations). The complete results are shown diagrammatically in Fig. 3. Each of the 84 dots represents the level of replication of a restriction fragment and its position on the walk; an artist's view

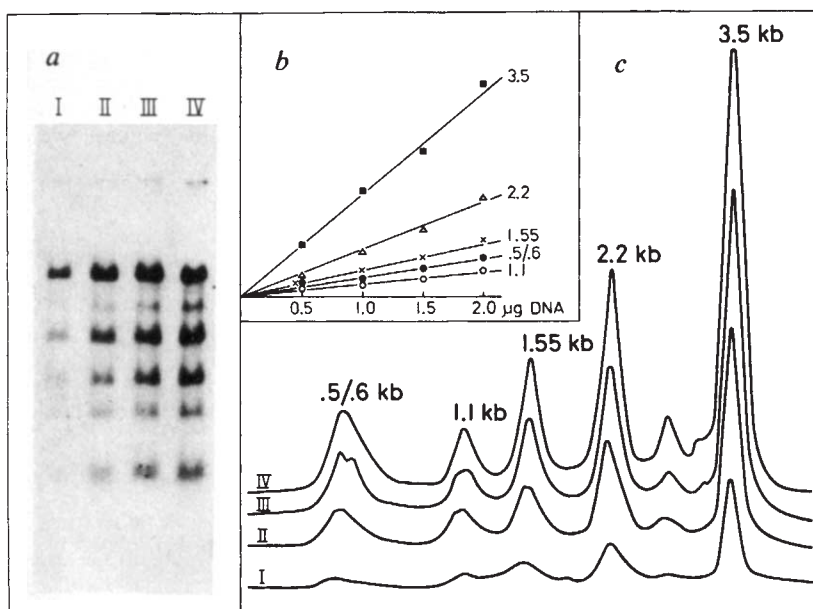


Fig. 1 Densitometry of autoradiographs of genomic Southern blot hybridizations does measure single copy restriction fragments. *a*, 0.5 µg (I), 1.0 µg (II), 1.5 µg (III) and 2.0 µg (IV) of *Drosophila* embryonic DNA brought to 2.0 µg with salmon sperm DNA were digested with *Eco*RI, size separated on 0.9% agarose gels and transferred to nitrocellulose. Filters were hybridized to the nick-translated recombinant 2160⁶, which comprises only single copy sequences, and autoradiographed. *b*, Plot of the heights of peaks [cm] obtained by scanning microdensitometry (*c*) of autoradiographs in *a*, as a function of *Drosophila* DNA concentration.

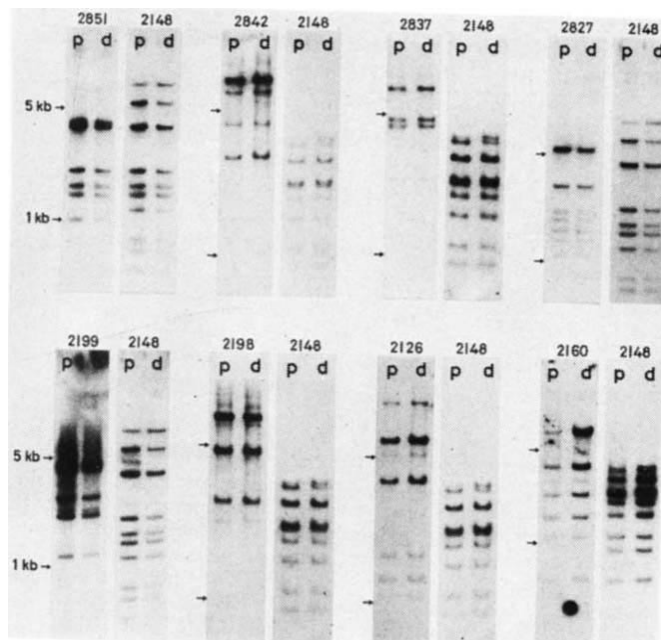


Fig. 2 Autoradiographs of genomic Southern blot hybridizations of polytene and diploid DNA with different segments of the walk. Polytene DNA was prepared from hand-dissected salivary glands of third instar larvae while diploid DNA was from embryos (0–16-h old). DNAs were digested with *Eco*RI (2842, 2837, 2198, 2126, 2160) or with *Bam*HI and *Hind*III (2851, 2827, 2199). Polytene (p) and diploid (d) DNAs were size-separated on parallel tracks of agarose gels and transferred to nitrocellulose. Filters were hybridized firstly to 2148⁶ for calibration and then to other probes whose positions on the walk are indicated in Fig. 3. The arrows indicate the 5-kb size (top) and 1-kb size (bottom). For each pair of panels, the panel on the right shows the calibration hybridization with 2148⁶ (equivalent to 2146 in Fig. 3); the panel on the left is the hybridization of the same filter to one of the segments of the walk. Filters were recycled by soaking in 0.1 M NaOH for 3 min at room temperature. The importance of this calibration is visible in the 2851 measurement: the signal is fainter in the diploid track to the same extent in the 2851 and in the 2148 hybridizations. Intensities and sizes of bands are as expected for unique sequences. Fainter bands are seen when the probe overlaps only partially with the fragment (5-kb and 11-kb fragments in 2126, for example) or in a few cases because of partial digests. Partial digestion, however, represents always less than 10% of the full digest (as seen on scans) and, hence, does not contribute significantly to the results. Heavier than expected bands are due to superposition of fragments of the same size (bands of about 4.5 kb in 2851, and of about 2.3 kb in 2148 digested by *Eco*RI for example).

of the polytene chromosome banding pattern is aligned with the DNA sequence map. The mean of the ratios for the 81 unique fragments is 0.97. The values range from 0.55 to 1.56 with a standard deviation of 0.25.

We see no local change in the level of polytenization, for no value goes beyond the 50% over or under-replication which we would consider significant. The large band 87E1,2 extends over 160 kb in the central part of the walk⁷. The relative abundance of about 35 fragments was measured in this interval, thus showing that there is no significant variation of level of DNA replication across this large band. The left part of the walk covers 9 ± 1 bands and interbands in 85 kb. The resolution at which we can place bands and interbands on the DNA map in this region does not allow us to tell whether particular restriction fragments are made of band or interband DNA⁷. However, cytological measurements of interband length led Beermann⁵ to estimate that their average DNA sequence content is about 2 kb. Therefore, the 9 ± 1 interbands of the region spans about 18 out of the 85 kb of sequence and should be largely represented in the set of 26 fragments (mean size 2.6 kb) which were measured. We conclude that there is no large variation of polyteny among

bands and interbands, and within a large band in this interval. Finally replication models in which bands and interbands have different levels of replication require that dormant replication forks are aligned at the band/interband boundaries. This should change the appearance of restriction fragments on genomic Southern blots. The genomic restriction pattern, however, was identical in polytene and diploid DNA all along the walk (not shown except the eight examples in Fig. 2).

Our study of a typical segment of the chromosome comprising bands of all sizes suggest that monotonous polyteny along the euchromatic arms of polytene chromosomes is a general feature. This is supported by the recent observation that a set of cloned sequences arising from different chromosomal locations have a similar level of replication in polytene chromosomes¹¹. Chromosomal constrictions may, however, be the unavoidable exceptions to the rule. Lyfschitz¹¹ found that the cluster of histone genes close to or within the constriction at 39D is under-replicated. We have determined that the bithorax complex of homeotic genes¹² which spans the 89DE constriction is also under-replicated in polytene chromosomes (P.S., A.S. and W. Bender, unpublished results).

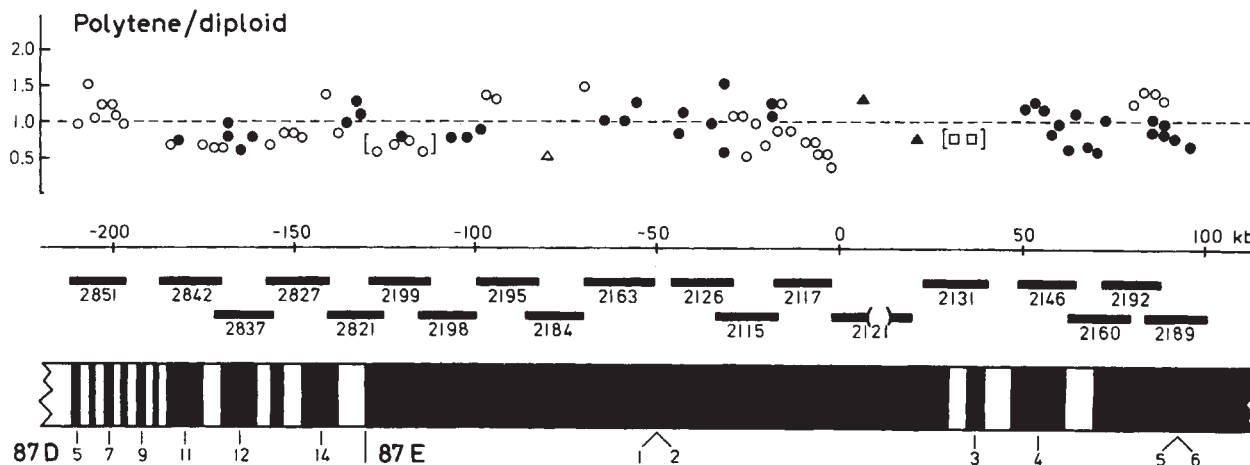


Fig. 3 Relative level of replication in polytene chromosomes of 84 restriction fragments spanning 13 ± 1 bands and interbands. Hybridizations were as in Fig. 2. The position of the probes is indicated with solid bars, and their numbers refer to ref. 6. Polytene/diploid ratios of individual restriction fragments are plotted as a function of their DNA map position. Calibration steps are explained in the text. The banding pattern is schematically aligned with the DNA map as described in ref. 7. ○, Fragments resulting from *Bam*HI and *Hind*III combined digestion; ●, fragments from *Eco*RI digestion; △, ▲, fragments larger than 5 kb, digested with *Bam*HI/*Hind*III or *Eco*RI respectively, when no smaller fragment could be used; □, fragments from *Bam*HI/*Hind*III which may contain moderately repeated sequences (not shown). Order of fragments in parentheses has not been determined. When two fragments co-migrated in gels, they were plotted at their respective positions on the walk with a ratio equal to the ratio of the sum of the signals.

We thank Professor Alfred Tissieres for his interest and support, and we acknowledge the help of Mrs M. Vollenweider and Mrs C. H. Tonka in the task of hand-dissecting salivary glands. This work was supported by a grant from the Swiss National Science Foundation.

Received 17 June; accepted 19 October 1983.

1. Painter, T. S. *Genetics* **19**, 179–188 (1934).
2. Sorsa, V. *Hereditas* **78**, 298–302 (1974).
3. Laird, C. D. *Cell* **22**, 869–874 (1980).
4. Laird, C. D., Wilkinson, L., Johnson, D. & Sandström, C. *Chromosomes Today* Vol. 7 (eds Bennet, M., Bobrov, M. & Hewitt, G.) 74–83 (George, Allen & Unwin, Hemel Hempstead, 1981).
5. Beermann, W. (ed.) *Results and Problems in Cell Differentiation* Vol. 4, 1–33 (Springer, New York, 1972).
6. Bender, W., Spierer, P. & Hogness, D. S. *J. molec. Biol.* **168**, 17–33 (1983).
7. Spierer, P., Spierer, A., Bender, W. & Hogness, D. S. *J. molec. Biol.* **168**, 35–50 (1983).
8. DuPraw, E. J. & Rae, P. M. *Nature* **212**, 598–600 (1966).
9. Lis, J. T., Prestidge, L. & Hogness, D. S. *Cell* **14**, 901–919 (1978).
10. Endow, S. A. & Glover, D. M. *Cell* **17**, 597–605 (1979).
11. Lytschitz, E. *J. molec. Biol.* **164**, 17–34 (1983).
12. Bender, W. *et al. Science* **221**, 23–29 (1983).
13. Hilliker, A. J., Clark, S. H., Chovnick, A. & Gelbart, W. M. *Genetics* **95**, 95–110 (1980).

Human hepatitis B vaccine from recombinant yeast

William J. McAleer, Eugene B. Buynak,
Robert Z. Maigetter, D. Eugene Wampler,
William J. Miller & Maurice R. Hilleman*

Division of Virus and Cell Biology Research, Merck Institute for Therapeutic Research, Merck Sharp & Dohme Research Laboratories, West Point, Pennsylvania 19486, USA

The worldwide importance of human hepatitis B virus infection and the toll it takes in chronic liver disease, cirrhosis and hepatocarcinoma, make it imperative that a vaccine be developed for worldwide application¹. Human hepatitis B vaccines^{2–6} are presently prepared using hepatitis B surface antigen (HBsAg) that is purified from the plasma of human carriers of hepatitis B virus infection. The preparation of hepatitis B vaccine from a human source is restricted by the available supply of infected human plasma and by the need to apply stringent processes that purify the antigen and render it free of infectious hepatitis B virus and other possible living agents that might be present in the plasma. Joint efforts between our laboratories and those of Drs W. Rutter and B. Hall led to the preparation of vectors carrying the DNA sequence^{7,8} for HBsAg and antigen expression in the yeast *Saccharomyces cerevisiae*⁹. Here we describe the development of hepatitis B vaccine of yeast cell origin. HBsAg of subtype adw was produced in recombinant yeast cell culture, and the purified antigen in alum formulation stimulated production of antibody in mice, grivet monkeys and chimpanzees. Vaccinated chimpanzees were totally protected when challenged intravenously with either homologous or heterologous subtype adr and ayw virus of human serum source. This is the first example of a vaccine produced from recombinant cells which is effective against a human viral infection.

Several alternative approaches to a hepatitis B vaccine are being developed. HBsAg has been expressed by several transformed mammalian cell lines, such as the human hepatoma line, PLC/PRF/5 (refs 10, 11), simian virus 40-infected monkey kidney cells¹² and mouse L cells¹³. These sources are of some concern, however, because the cell lines may be neoplastic. Although HBsAg has been cloned in bacteria^{7,14}, expression was very weak. Other laboratories^{15–19} have described the synthesis of oligopeptides that carry antigenic determinants of HBsAg but their potency in animals is low and much work will need to be done to potentiate antigenicity. Smith and collaborators²⁰ have described the construction of a recombinant

Table 1 Antigenic potency in mice of HBsAg purified from yeast and from human plasma

Vaccine source	Antigen dose per injection (µg protein)	Anti-HBsAg response after vaccination	
		No. positive/total	GMT
Human plasma (lot 799-2)	10	9/10	563
	2.5	10/10	2,235
	0.625	4/9	32
	0.156	0/10	4
	ED ₅₀ 0.639		
Yeast (lot 81-4)	40	10/10	5,432
	10	10/10	3,400
	2.5	8/10	673
	0.625	8/10	967
	ED ₅₀ <0.625		

Groups of 10 5-week-old ICR/Ha mice propagated in our laboratories were given a single 1-ml injection intraperitoneally of serial fourfold dilutions of yeast or human plasma vaccine in alum diluent. The mice were bled individually and tested for serum antibody level 4 weeks later. Human plasma vaccine, lot 799-2, was prepared in these laboratories^{2–4}. Yeast-derived vaccine, lot 81-4, was purified as described in Fig. 1 legend and adsorbed to alum. GMT, geometric mean titre, expressed in AUSAB units; ED₅₀, dose required to seroconvert 50% of the mice.

vaccinia virus which expresses HBsAg and have proposed its use as a live attenuated vaccine; its antigenic potency has been demonstrated but whether such a vaccine would be safe and effective in man is still unknown.

Valenzuela *et al.*⁹ originally reported that yeast cells are able not only to express the HBsAg gene but also to assemble the polypeptides into particles that have much the same appearance as particles isolated from human plasma and which are immunogenic in mice. Since then, other laboratories^{21,22} have shown that HBsAg produced in yeast is antigenic in rabbits and guinea pigs. With such progress, recombinant yeast has become an attractive alternative to human plasma as a source of antigen for hepatitis B vaccine.

For vaccine preparation, the HBsAg used was of subtype adw and was produced in fermentation cultures of *S. cerevisiae* carrying an expression vector using yeast alcohol dehydrogenase I as a promoter. The yeast strain used in these studies was obtained from G. Ammerer (University of Washington) and is similar to the strain described by Valenzuela *et al.*⁹ in which the production of HBsAg in yeast was first reported.

Cells were collected by centrifugation and broken by homogenization with glass beads²³. HBsAg particles were purified from the clarified extract by immune affinity chromatography using goat antibody to human HBsAg. Electron microscopy (Fig. 1) revealed a homogeneous array of particles free of extraneous morphological entities. The UV absorption pattern was the same as for the plasma antigen, with an $E^{1\%}$ of 45. SDS-polyacrylamide gel electrophoresis (Fig. 2) in reducing conditions revealed a major band at molecular weight 23,000 (23K) corresponding to the non-glycosylated polypeptide which is the major polypeptide of the viral envelope. In this respect it differs from the plasma antigen which has, in addition to the 23K polypeptide, a glycosylated derivative which migrates at 27K. The yeast and plasma antigens differ also in their reactivity in the radioimmunoassay (RIA) (AUSRIA II, Abbott). RIA reactivity of purified yeast-derived HBsAg varied from preparation to preparation in the range 20–50% of the reference human antigen.

Because of this reduced radioimmune reactivity, and because the yeast antigen is not glycosylated, it was important to determine whether the antigen was immunogenic. To test both antigenicity and immunogenicity in animals, purified antigen was formulated into a vaccine by adsorbing on alum adjuvant to contain 40 µg HBsAg protein and 0.5 mg aluminium (hydroxide) per 1 ml dose.

* To whom reprint requests should be addressed.

Table 2 Antigenic potency in grivet monkeys of HBsAg purified from yeast and from human plasma

Vaccine source	Antigen dose per injection (μg protein)	Week 4	Anti-HBsAg response after initial vaccine dose (geometric mean titre)		
			Week 8	Week 12	Week 52
Human plasma (lot 86016)	10	36	213	170	127
	2.5	343	6,227	17,348	9,924
	0.625	53	4,642	3,164	5,688
	0.156	15	128	83	358
Yeast (lot 81-4)	40	88	1,078	7,103	11,554
	10	184	877	8,489	4,984
	2.5	225	1,168	6,361	10,868
	0.625	109	925	518	313

A group of four initially seronegative grivet monkeys (*Cercopithecus aethiops*), weighing 3–5 kg, were each given two 1-ml intramuscular (i.m.) doses of yeast or human plasma vaccine 4 weeks apart. Dilutions of antigen were made in alum placebo of the same composition as the vaccine. Animals were bled at biweekly intervals for 1 yr and tested for antibody to HBsAg by using a commercial RIA kit (AUSAB, Abbott). Protein was measured by the method of Lowry²⁹. Human plasma lot 86016 was prepared in these laboratories²⁻⁴.

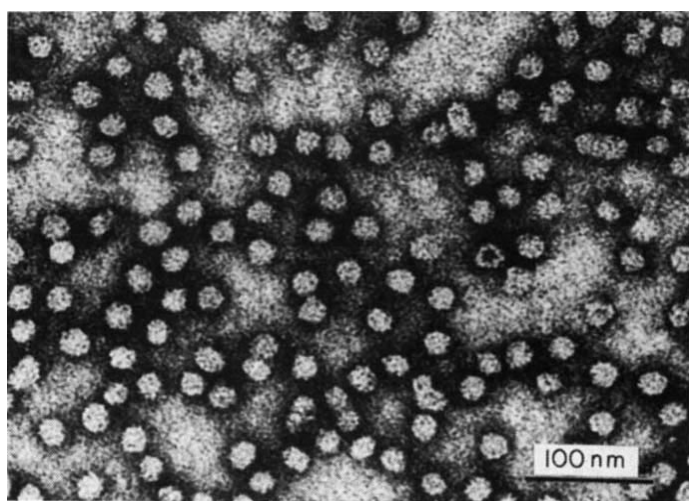


Fig. 1 Electron micrograph of HBsAg particles from recombinant yeast. Cells were grown in a 335-l fermentation vessel, collected by centrifugation, resuspended in an equal volume of 0.01 M sodium phosphate pH 7.5, containing 0.01% Triton X-100, and disrupted by rapid stirring with glass beads in a Dyno-Mill (Impandex; see ref. 23). The resulting extract was clarified by centrifugation for 90 min at 10,000g. The clarified yeast extract was applied to a column of Sepharose 4B to which had been attached goat antibody to human HBsAg. The column was developed at a flow rate of 2 column vol per h. Extraneous protein was washed away with 5 column vol of buffer A and the HBsAg was eluted with 3 M NH_4SCN . Fractions containing HBsAg were pooled and thiocyanate was removed by dialysis against 0.01 M sodium phosphate pH 6.8, containing 0.15 M NaCl. Dialysed antigen was diluted to $40 \mu\text{g ml}^{-1}$ and visualized by negative staining with 2% phosphotungstic acid.

Studies in mice (Table 1) showed the yeast-derived antigen to be at least as antigenic as the antigen purified from human plasma. Grivet monkeys also developed antibody following vaccination with the yeast-derived antigen (Table 2). A single injection of the vaccine at all dose levels resulted in seroconversion of all the animals in both vaccine groups. These results were important as they showed that high antibody titres were maintained for at least a year.

Protective efficacy was tested for by using susceptible chimpanzees. The four chimpanzees that received the recombinant vaccine developed antibody in substantial titre following vaccination (Table 3). Following challenge with infectious human plasma, all four vaccinated animals were protected. By contrast, all four unvaccinated animals developed hepatitis B virus infection with positive antigenaemia, antibody to hepatitis B core antigen (anti-HBcAg), elevation of serum glutamic oxalacetic

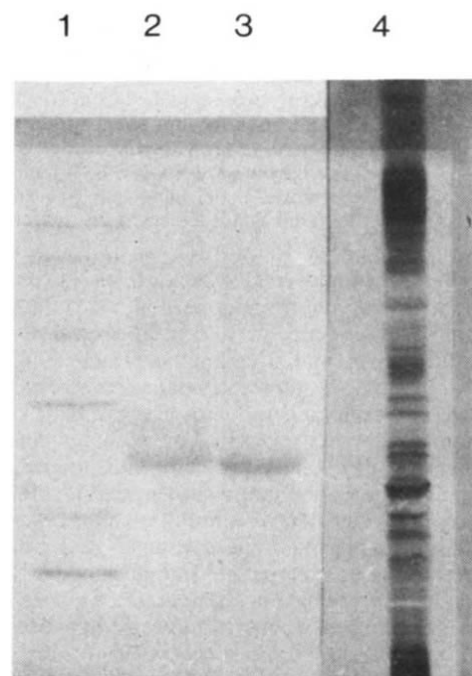


Fig. 2 SDS-polyacrylamide gel electrophoresis of cell culture and yeast-derived HBsAg. All samples were reduced, denatured and electrophoresed as described by Laemmli³⁰. After electrophoresis, polypeptides were visualized with Coomassie brilliant blue (lanes 1–3) or with the silver stain procedure described by Morrissey³¹ (lane 4). Lane 1, molecular weight standards (3 μg each): phosphorylase b (94K), bovine serum albumin (68K), ovalbumin (43K), carbonic anhydrase (30K), soybean trypsin inhibitor (21K) and lysozyme (14.3K). Lane 2, 30 μg of HBsAg from the human hepatoma cell line PLC/PRF/5 (ref. 10), also purified from yeast as described in Fig. 1 legend. Lane 3, 30 μg of HBsAg purified from yeast as described in Fig. 1. Lane 4, 10 μg of clarified yeast extract as described in Fig. 1 legend.

transaminase (SGOT) and serum glutamic pyruvic transaminase (SGPT), and liver histopathology. It is important to note that the animals were protected against both subtype adr and ayw challenge even though the vaccine is of the adw subtype.

Yeast fermentation technology is well established and we have shown that HBsAg can be isolated from yeast extracts in a highly purified form by a single application of immune affinity chromatography. Vaccine made from this antigen is equally as potent as human plasma-derived vaccine in stimulating antibodies in mice, and is protective in challenge experiments in chimpanzees. Antibodies raised by yeast-derived vaccine persisted for at least a year in monkeys, showing no important deviation from that of the plasma vaccine.

Table 3 Protective efficacy in chimpanzees of HBsAg purified from yeast and from human plasma

Injection	Chimp no.	Before challenge		After challenge (week of onset or weeks of duration)								Liver pathology onset
		Anti-HBsAg titre (at 12 weeks)	Antigen subtype	HBsAg		Anti-HBcAg		SGOT elevation		SGPT elevation		
				Onset	Duration	Onset	Duration	Onset	Duration	Onset	Duration	
Yeast vaccine (lot 81-4)	110	1,830	adr	—	—	—	—	—	—	—	—	—
	138	540	adr	—	—	—	—	—	—	—	—	—
	103	18,300	ayw	—	—	—	—	—	—	—	—	—
	120	7,200	ayw	—	—	—	—	—	—	—	—	—
Unvaccinated controls	111	<8	adr	10	10	15	9	17	3	17	6	20
	128	<8	adr	8	11	12	12	17	3	16	5	20
	127	<8	ayw	6	14	12	12	13	3	13	7	16
	130	<8	ayw	6	18	10	14	22	1	14	10	24

Eight chimpanzees, each weighing 40–60 kg, were selected for study based on negative findings in tests for HBsAg, anti-HBsAg, anti-HBcAg, elevation in transaminase, liver histopathology and tuberculin reaction. The animals were separated into two groups, four test animals and four controls. Each of the four test animals was given three 40- μ g doses of yeast-derived HBsAg vaccine in 1 ml volume i.m. at 4-week intervals. All eight animals were then challenged by intravenous injection of 1,000 chimpanzee infectious doses of subtype adr or ayw virus in 1 ml of human hepatitis B plasma. Antigen and antibody titres were measured by commercial (Abbott) RIA kits (AUSRIA, AUSAB and CORAB for HBsAg, anti-HBsAg and anti-HBcAg, respectively). SGOT and SGPT assays were performed by the Sigma-Frankel (no. 505) and by the UV absorption (Boehringer-Mannheim) procedures, respectively. SGOT titres >40 and SGPT titres >30 were considered elevated. The subtype adr and ayw human plasmas used for challenge were obtained from Drs R. Gerety and E. Tabor of the Office of Biologics, US Food and Drug Administration; they were of measured viral infectiousness for chimpanzees and were subtyped serologically. The animals were bled at weekly intervals during the 36-week period of observation, covering 12 weeks before virus challenge and 24 weeks after. Liver biopsies were taken at 4-week intervals using a Menghini 16T needle. The tissues were fixed in 10% buffered formalin solution and the haematoxylin/eosin-stained sections were prepared by Dr A. Phelps of these laboratories under blind code number. The tests were carried out in animals that were held in isolation in the facilities of Dr William E. Greer at the Gulf South Research Institute, New Iberia, Louisiana. —, All remained negative.

Human HBsAg is composed of a sequence of 226 amino acids of which the a antigen determinant is dominant. Small differences in amino acid sequence may occur at several positions in the polypeptide chain and are responsible for the subtype specificities²⁴. In previous studies, chimpanzees that were cross-challenged with heterologous subtypes of hepatitis B virus after recovery from infection or vaccination with human plasma-derived antigens, were solidly protected due to the common group specificity of the dominant a antigen that is present in all HBsAg subtypes²⁵. A protective efficacy trial in man of subtype ad vaccine of human plasma origin has shown strong protection against the homologous subtype^{26,27} and, most recently, against the heterologous subtype ay²⁸ in studies carried out on the staffs of renal dialysis centres where subtype ay hepatitis is most common. The positive cross-protection afforded against heterologous subtype ayw virus challenge in chimpanzee immunized with type adw vaccine of yeast origin, indicates that the a antigen remains dominant in the recombinant-produced antigen obtained from human plasma.

We thank Dr C. E. Carty and F. X. Kovach for assistance in fermentation, B. J. Harder, N. B. Grason and J. Bailey for assistance in antigen purification, Dr B. Wolanski and R. Ziegler for electron microscopy, and H. E. Darmofal, J. T. Deviney, K. I. Guckert, R. R. Roehm and L. W. Stanton for assistance in the animal tests.

- Dubois, M. F., Pourcel, C., Rousset, S., Chany, C. & Tiollais, P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4549–4553 (1980).
- Burrell, J. C., Mackay, P., Greenway, P. J., Hofschneider, P. H. & Murray, K. *Nature* **279**, 43–47 (1979).
- Lerner, R. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 3403–3407 (1981).
- Dreesman, G. R. *et al. Nature* **295**, 158–160 (1982).
- Bhatnagar, P. K. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 4400–4404 (1982).
- Prince, A. M., Ikram, H. & Hopp, T. P. *Proc. natn. Acad. Sci. U.S.A.* **79**, 579–582 (1982).
- Gerin, J. L. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 2365–2369 (1983).
- Smith, G. L., Mackett, M. & Moss, B. *Nature* **302**, 490–495 (1983).
- Miyahara, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 1–5 (1983).
- Hitzeman, R. A. *et al. Nucleic Acids Res.* **11**, 2745–2763 (1983).
- Deters, D., Muller, U. & Homberger, H. *Analyt. Biochem.* **70**, 263–267 (1976).
- Ono, Y. *et al. Nucleic Acids Res.* **11**, 1747–1757 (1983).
- Gerety, R. J., Tabor, E., Purcell, R. H. & Tyeryar, F. J. *J. infect. Dis.* **140**, 642–648 (1979).
- Szmuness, W., Stevens, C. E., Zang, E. A., Harley, E. J. & Kellner, A. *Hepatology* **1**, 377–385 (1981).
- Francis, D. P. *et al. Ann. intern. Med.* **97**, 362–366 (1982).
- Szmuness, W. *et al. New Engl. J. Med.* **307**, 1481–1486 (1982).
- Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *biol. Chem.* **193**, 265–275 (1951).
- Laemmli, U. K. *Nature* **227**, 680–685 (1970).
- Morrissey, J. H. *Analyt. Biochem.* **117**, 307–310 (1981).

Antisera to a synthetic peptide of the *sis* viral oncogene product recognize human platelet-derived growth factor

Henry L. Niman

Department of Molecular Biology, Research Institute of Scripps Clinic, La Jolla, California 92037, USA

It has recently been reported that the sequences of the *sis* oncogene of simian sarcoma virus (SSV) and of human platelet-derived growth factor (PDGF) are very similar^{1,2}, establishing the most solid link yet between the mitogenic actions of growth factors and the transforming proteins of retroviruses. To investigate molecular mechanisms of transformation I have produced antisera against synthetic peptides corresponding to segments of the protein sequences predicted by the nucleotide sequences of viral oncogenes. Applying this approach to the case of *sis* and PDGF, I report here the results of probing outdated human platelets with an antiserum directed against a synthetic peptide representing residues 139–155 of the predicted sequence of the SSV transforming protein, p28^{sis} (ref. 3). I detected peptides of apparent molecular weights (MWs)

Received 11 July; accepted 16 November 1983.

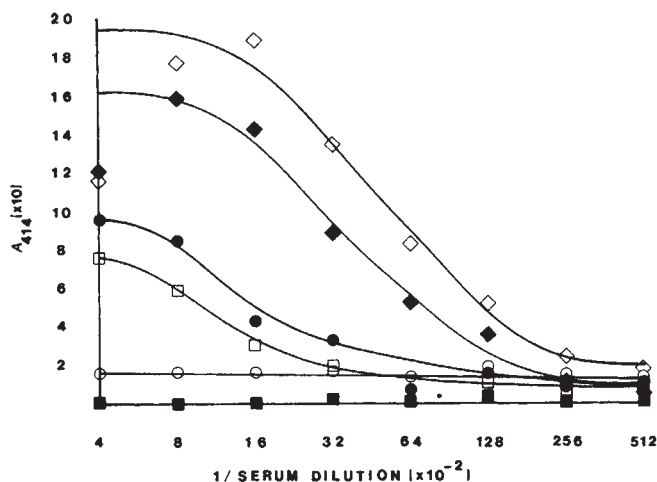
- Deinhardt, F. & Gust, I. D. *Bull. Wild Hlth Org.* **60**, 661–691 (1982).
- Hilleman, M. R. *et al. Viral Hepatitis* (eds Vyas, G. N., Cohen, S. N. & Schmid, R.) 525–537 (The Franklin Institute Press, Philadelphia, 1978).
- Hilleman, M. R. *et al. in Viral Hepatitis 1981 int. Symp.* (eds Szmuness, W., Alter, H. & Maynard, J.) 385–397 (The Franklin Institute Press, Philadelphia, 1982).
- Bynack, E. B. *et al. J. Am. med. Ass.* **235**, 2832–2834 (1976).
- Adamowicz, P. *et al. in INSERM Symp. No. 18* (eds Maupas, P. & Guesry, P.) 37–49 (Elsevier, Amsterdam, 1981).
- Coutinho, R. A. *et al. Br. med. J.* **286**, 1305–1308 (1983).
- Valenzuela, P. *et al. Nature* **280**, 815–819 (1979).
- Edman, J. C., Hallowell, R. A., Valenzuela, P., Goodman, H. M. & Rutter, W. J. *Nature* **291**, 503–506 (1981).
- Valenzuela, P., Medina, A., Rutter, W. J., Ammerer, G. & Hall, B. D. *Nature* **298**, 347–350 (1982).
- Alexander, J. J., Bey, E. M., Geddes, E. W. & Lecatsas, G. S. *Afr. med. J.* **50**, 2124–2128 (1976).
- Barin, F., Maupas, P., Coursaget, P., Goudeau, A. & Chiron, J. P. *in INSERM Symp. No. 18* (eds Maupas, P. & Guesry, P.) 263–266 (Elsevier, Amsterdam, 1981).
- Moriarty, A. M., Hoyer, B. H., Shih, J. W.-K., Gerin, J. L. & Hamer, D. H. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2606–2610 (1981).

The activities of two sequence-specific antisera are shown in Fig. 1. These antisera were prepared by immunizing mice with synthetic peptides coupled to keyhole limpet haemocyanin (KLH). The anti-*sis* antiserum was made against the peptide shown in Table 1. The anti-*myb* serum was made against a peptide representing residues 2-18 of the predicted sequence of the transforming protein of avian myeloblastosis virus². The anti-*sis* serum reacts with the *sis* synthetic peptide and KLH but not with the *myb* peptide. The anti-*myb* antiserum reacts with the *myb* peptide and KLH but not the *sis* peptide. To test these sera for activity with cognate structures in intact molecules, purified PDGF (given by R. Ross) and an extract of marmoset cells nonproductively infected with SSV (NPV-SSV) were probed with the *sis* and *myb* antisera. Probing the purified

Genotype	Accession	Sequence
p28 ^{Sis}	1	MTLTWQGDPIPEELYKMLSGHSIRSFDDLQRLQLQGDSGKE
p28 ^{Sis}	41	DGAELDLNMRSHSGGELES LARGKRSLGSLSVAE PAMIA
PDGF-2 ⁺	1	SLGSLTIAEPAMIA
PDGF-2 ⁺	1	SLGSLTIAEPAMIA
PDGF-1 [*]	1	SIEEAVPA
PDGF-1 ⁺	1	SIEEAVPA
p28 ^{Sis}	81	<u>E</u> CKTRTEVFEISRRLLIDRTNANFLVWPPECVEVQRCSGCCN
PDGF-2 [*]	15	ECKTREVEVFCICRRL?DR????????PPCDEVKRCCTGCCN
PDGF-2 ⁺	15	ECKTRTEVFEISRRLLID?TTANFLVWPPECVEVQRCSGCCN
PDGF-1 [*]	9	VCKTRIVIEYEISRRELD???ANFL
PDGF-1 ⁺	9	VCKTRIVIEYEIPRSQVDPTSANFLVWPPECVEVQRCSGCCN
p28 ^{Sis}	121	NRNVQCRPTQVQLRPVQVRKIEIVRKKPIFKKATVTLEDH
PDGF-2 [*]	55	NRNVKCRPSQVQLRP?QVRKIEIVRK
PDGF-2 ⁺	55	NRNVQCRPTQVQLRPVQVRKIE????KKPIFKKA?V?LEDH
PDGF-1 ⁺	49	NRNVQCRPTQVQL?PVQ????????KKPIFKKA?V?LEDH
p28 ^{Sis}	161	LACKCEIVAAARAVTRS PGTSQEQRAKTTQSRVTIIRTVR
PDGF-2 ⁺	95	LACKC?IVAAA
PDGF-1 ⁺	89	LACKC?IVAAA
p28 ^{Sis}	201	RRPPKGKHKRCKKHTHDKTALKETLGA

* Ref. 1; † ref. 2.

To analyse the size of PDGF in platelets, an extract made from outdated platelets was probed with the anti-*sis* and anti-*myb* antisera. The platelets were treated as described for the first two steps of PDGF purification⁴. This extract was electrophoresed into a polyacrylamide gel and then transferred to nitrocellulose¹¹. The extract was then probed with the anti-*sis*



Methods: Peptides representing residues 139–155 of the predicted sequence of the SSV transforming protein p28^{sis} (ref. 3) or residues 2–18 of the predicted sequence of the avian myeloblastosis virus⁹ were synthesized using solid state methods as described elsewhere³¹. A cysteine was added to the carboxyl end of each sequence for coupling purposes. The amino acid composition of each peptide was determined. For immunization, the peptides were coupled to KLH as described elsewhere³². The peptide conjugates were administered at ~50 µg per mouse per injection. On day 0, each conjugate was mixed with an equal volume of Freund's complete adjuvant and injected intraperitoneally (i.p.). On day 19 each conjugate was mixed with a final concentration of 5 mg ml⁻¹ alum and injected i.p. For the *sis* conjugate, the final injection, in phosphate-buffered saline (PBS), was given intravenously (i.v.) on day 62 and plasma was taken from blood on day 67 by orbital puncture. For the *myb* conjugate a second alum injection was given on day 41; the final injection was given i.v. on day 143 and the plasma was collected on day 148. To test the specificity of the antisera, 250 ng of unconjugated peptide or 500 ng of KLH were dried onto the bottom of microtitre wells and fixed with methanol as described previously³³. To prevent nonspecific adsorption, the wells were incubated at 37°C for 4 h with 3% bovine serum albumin (BSA). 25 µl of serial twofold dilutions of mouse plasma, starting with a 1:400 dilution in tissue culture medium supplemented with 10% fetal calf serum, were incubated with peptide or KLH for 16 h at 25°C. After washing 10 times with distilled H₂O, 25 µl of rabbit anti-mouse κ (Litton) diluted 1:500 with 1% BSA in PBS were added for 2 h at 37°C. Then after washing 10 times with distilled H₂O, 25 µl of glucose oxidase-conjugated goat anti-rabbit IgG diluted 1:500 in 1% BSA in PBS were added for 1 h at 37°C. The amount of glucose oxidase bound was determined by using 50 µl of 100 µg ml⁻¹ ABTS dye (Boehringer-Mannheim) in the presence of 1.2% glucose, 10 µg ml⁻¹ horseradish peroxidase in 0.1 M phosphate buffer, pH 6.0. The optical density at 414 nm was read in a Titertech microscanner.

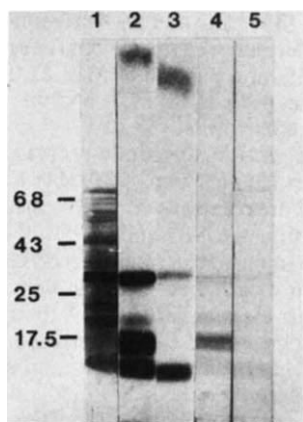


Fig. 2 Immunological detection of purified PDGF and SSV transforming protein on nitrocellulose filters. Approximately 700 ng of purified PDGF with 150 μ g carrier BSA (lanes 1, 3) or a lysate of 10^6 SSV-infected cells prepared as described previously¹² (lanes 2, 4) were electrophoresed onto a 7.5–17% polyacrylamide gel and transferred to nitrocellulose as described¹¹. After blocking nonspecific binding with 3% BSA and 0.1% Triton X-100 in PBS, 4 ml of mouse anti-*sis* or anti-*myb* sera diluted 1:100 with 3% BSA, 0.1% Triton X-100 were incubated for 16 h at 25 °C. After washing three times with 0.1% Triton X-100 in PBS, 10 ml of rabbit anti-mouse IgG1 (Litton) diluted 1:300 with 3% BSA, 0.1% Triton X-100 in PBS were incubated with the nitrocellulose for 1 h at 37 °C; after washing three times with 0.1% Triton X-100 in PBS, the strips were incubated with 10^7 c.p.m. of 125 I-protein A and washed with 0.1% Triton X-100 in PBS. The 125 I-labelled lanes were visualized with XRP-1 film (Kodak) and Cronex Hi-plus intensifying screen (Dupont) (–70 °C for 3.5 h). The anti-*sis* serum (lanes 1, 2) was obtained from a mouse immunized using the following protocol. On day 0 the *sis* peptide conjugate was given i.p. at ~50 μ g in Freund's complete adjuvant. On days 15 and 31, 50 μ g of the conjugate was mixed with 5 mg ml⁻¹ alum and injected i.p. The plasma was isolated from blood taken on day 77 by orbital puncture. The anti-*myb* serum (lanes 3, 4) was obtained as described in Fig. 1 legend. The molecular weight standards were lysozyme, soybean trypsin inhibitor, carbonic anhydrase, ovalbumin, BSA and phosphorylase *b*. All of the bands visualized in lane 4 with the anti-*myb* serum were also seen on longer exposures of lane 2 (probed with the anti-*sis* serum). Only a major band of 56K is seen specifically in lane 2. (Longer exposure revealed a specific 28K protein in lane 2; data not shown.)

and anti-*myb* antisera (Fig. 3): lane 1 is a stain of the total protein bound to nitrocellulose. Although many proteins could be visualized with amido-black staining, PDGF was not present in sufficient quantities to be seen. Figure 3, lanes 2 and 3 show the results of probing with the anti-*sis* and anti-*myb* antisera. Bound antibody was detected with 125 I-labelled protein A. The specific bands detected with the anti-*sis* antisera migrated with an apparent MW of ~30K as well as 18–16K. In addition, a minor band of ~21K was detected. Visualization of antibody by benzidine staining¹² of peroxidase-conjugated rabbit anti-mouse IgG (Fig. 3, lanes 4, 5) produced similar results except that the higher molecular weight proteins were resolved as a major band of ~31K and a minor band of 30K. Similarly, the lower bands were resolved as major bands of ~18K and 16K, with protein of a more heterogeneous size migrating between these two bands. In addition, a minor band of ~21K was detected. The specificity of the anti-*sis* antiserum was further tested by preincubation with synthetic peptide conjugates. Incubation of anti-*sis* antiserum with the unrelated *myb* peptide conjugate (Fig. 3, lane 1) did not affect binding to the specific bands seen in Fig. 2, whereas incubation with the *sis* peptide conjugate eliminated detectable binding to all of the specific bands. Only a nonspecific binding (also detected by the anti-*myb* serum, lane 3) which co-migrated with the major protein in the extract was visualized. The mobilities of ~31K and 17K correspond to the unreduced and reduced forms of human PDGF⁴⁻⁸. The detection of the ~31 K and 21K bands indicates that the

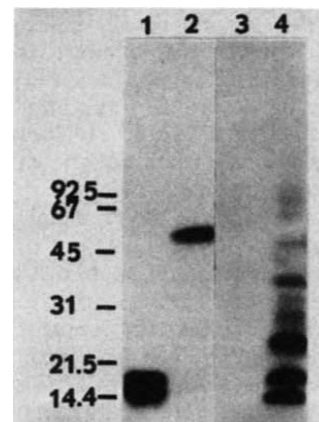


Fig. 3 Immunological detection of nonreduced and reduced PDGF on nitrocellulose filters. Lane 1 is total protein stained with amido-black. Lanes 2 and 3 were developed with 125 I; lanes 4 and 5 with dimethoxybenzidine. The platelet extract was probed with the anti-*sis* serum (lanes 2, 4) or the anti-*myb* serum (lanes 3, 5). The molecular weight standards were BSA, ovalbumin, chymotrypsinogen and β -lactoglobulin.

Methods: Using outdated human platelets (16 units; obtained from the San Diego Blood Bank), the first two steps of the PDGF purification were followed⁴. Platelets were collected by centrifugation at 28,000g for 20 min at 4 °C, then washed twice by resuspending them in 400 ml of 9-vol 17 mM Tris-HCl pH 7.4, 0.15 M NaCl, 0.1% glucose; 1 vol 0.8 g citric acid monohydrate, 2.2 g dextrose (anhydrous), 2.6 g sodium citrate (dihydrate) in 100 ml, and centrifuged at 28,000g for 10 min at 4 °C. The washed platelets were resuspended in 16 ml of 0.08 M NaCl, 0.01 M PO₄ pH 7.4 and boiled for 10 min. Phenylmethylsulphonyl fluoride and Trasylol were added to a final concentration of 1 mM and 3% respectively. The lysate was mixed with 8 ml CM Sephadex C-50 previously equilibrated with 0.08 M NaCl, 0.01 M PO₄ pH 7.4. The platelet pellet was extracted twice with 1 M NaCl for 24 h at 4 °C. After pelleting the insoluble material, the supernatant was dialysed against 0.08 M NaCl, 0.01 M PO₄ pH 7.4. PDGF extracted from the pellet was also mixed with 0.5 vol CM Sephadex C-50 equilibrated with 0.08 M NaCl, 0.01 M PO₄ pH 7.4. The beads were poured into a 15 $\times$ 1.5 cm column, and washed with 6 column vol of 0.08 M NaCl, 0.01 M PO₄. PDGF was eluted with two column vol of 1 M NaCl. Trasylol was added to a final concentration of 3%, then the eluate was dialysed against 0.08 M NaCl, 0.01 M PO₄ pH 7.4. All three extracts were pooled and concentrated with a 10K exclusion filter (Amicon). The extract from ~2.5 units of platelets was adjusted to 0.5% SDS, 5% 2-mercaptoethanol, boiled for 2 min and electrophoresed into a 5–17% polyacrylamide gel. The protein was electrophoretically transferred to nitrocellulose as described¹¹ and cut into strips. After blocking nonspecific binding with 3% BSA and 0.1% Triton X-100 in PBS, 4 ml of mouse plasma diluted 1:200 were incubated with the nitrocellulose strips. After washing three times with 0.1% Triton X-100 in PBS, the strips were incubated with either 10^7 c.p.m. 125 I-labelled protein A or a 1:1,000 dilution of peroxidase-conjugated goat anti-mouse serum (Tago), and washed with 0.1% Triton X-100 in PBS. The 125 I-labelled strips were visualized with XRP-1 film (Kodak) and Cronex Hi-plus intensifying screens (Dupont) (–70 °C for 48 h). The peroxidase conjugate was developed with 0.009% H₂O₂, 0.0025% 3,3'-dimethoxybenzidine dihydrochloride (Kodak), 10 mM Tris pH 7.4.

higher molecular weight forms of PDGF in platelets are resistant to reduction.

The data presented immunologically confirm the sequence homology between the SSV oncogene and human PDGF. The factor(s) responsible for the resistance of some of the PDGF and *sis* gene product to reduction is unknown. However, the mitogenic activity of PDGF is associated with the nonreduced forms^{6,13,14}. Stimulation of tyrosine-specific kinase activity by PDGF represents the third example of phosphorylation associated with proteins of oncogenes. Members of the *ras* gene family bind GTP and autophosphorylate pp21 (ref. 15). Members of the oncogene family having sequence homology with the carboxyl portion of pp60^{src} phosphorylate IgG *in vitro* and stimulate

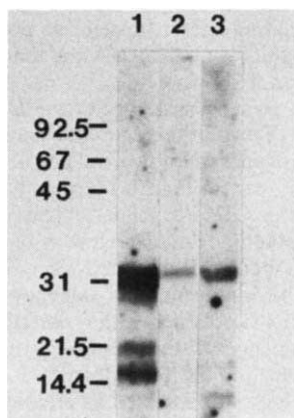


Fig. 4 Immunoreactivity with PDGF is blocked by the synthetic peptide. Platelets were extracted as described in Fig. 3 legend except that Trasylol was added to a final concentration of 1% after elution from the CM Sephadex C-50 column. The extract from ~2.5 units of platelets was adjusted to 0.5% SDS, 5% 2-mercaptoethanol, boiled for 5 min and electrophoresed into a 7.5–17% polyacrylamide gel. Transfer was as described in Fig. 3 legend. The plasma used in lanes 1 and 2 was taken from a mouse immunized as described in Fig. 2 legend except that there was no i.v. injection; instead, the third i.p. injection was in alum and given on day 215 and the plasma was collected on day 223 by orbital puncture. Lane 1, platelet extract probed with anti-*sis* antiserum preincubated with the *myb* conjugate (30 µg for 1 h at 37 °C). Lane 2, anti-*sis* antiserum preincubated with the *sis* conjugate (30 µg for 1 h at 37 °C). Lane 3, extract probed with the anti-*myb* serum. The molecular weight standards were lysozyme, soybean trypsin inhibitor, carbonic anhydrase, ovalbumin, BSA and phosphorylase *b*.

tyrosine phosphorylation in cell extracts^{16–23}. Although p28^{sis} does not seem to autophosphorylate or label IgG *in vitro*¹⁰, PDGF stimulation of tyrosine phosphorylation in cell extracts^{24–27} suggests that phosphorylation reactions may be involved in numerous pathways important in cell proliferation and transformation.

These pathways can be further characterized using sequence-specific antibodies. This report describes a peptide that readily produces antibodies which recognize the synthetic peptide as well as the cognate structure in the intact molecule. Such peptides can be used for high-frequency isolation of hybridomas secreting useful monoclonal antibodies¹²; these antibodies can be used to map the active site of PDGF. Antibodies which do not block activity can be used to isolate PDGF receptors using immunoaffinity procedures. Monoclonal antibodies will be useful in the characterization of PDGF-related molecules and associated substrates (for example, tumour cells of mesenchymal origin have been shown to express RNA related to *v-sis*²⁸). In addition, monoclonal antibodies against sequences unique to each of the PDGF peptides can facilitate their isolation and separation for additional sequencing. Apart from the use of sequence-specific monoclonal antibodies for the characterization of biochemical pathways used in cell proliferation and transformation, selected antibodies should have diagnostic and therapeutic value.

During the preparation of this manuscript, Robbins *et al.* reported²⁹ the detection of a 56K dimer of the *sis* transforming gene product and the immunological detection of this gene product by an antiserum directed against PDGF. Similarly, Deuel *et al.*³⁰ showed the reactivity of an anti-PDGF antiserum with a 20K protein in SSV-infected cells.

I thank R. Ross, D. Bowen-Pope and E. Raines for supplying PDGF, R. A. Lerner for discussions, R. A. Houghten for synthesizing and analysing peptides, the San Diego Blood Bank for supplying outdated platelets, R. Naso for supplying SSV-infected marmoset cells, and B. Heilman and D. Schloeder for technical assistance. This work was supported in part by NIH grant CA25803. This is publication no. 3162-MB from the Research Institute of Scripps Clinic.

Received 12 August; accepted 18 November 1983.

1. Doolittle, R. F. *et al. Science* **220**, 275–277 (1983).
2. Waterfield, M. D. *et al. Nature* **304**, 35–39 (1983).
3. Devare, S. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 731–735 (1983).
4. Antonides, H. N., Scher, C. D. & Stiles, C. D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1809–1813 (1979).
5. Heldin, C. H., Westermark, B. & Wasteson, A. *Biochem. J.* **193**, 907–913 (1981).
6. Antonides, H. N. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7314–7318 (1981).
7. Deuel, T. F. *et al. J. biol. Chem.* **256**, 8896–8899 (1981).
8. Johnson, A., Heldin, C. H., Westermark, B. & Watson, A. *Biochem. biophys. Res. Commun.* **104**, 66–74 (1982).
9. Rushlow, K. E. *et al. Science* **216**, 1421–1423 (1982).
10. Robbins, K. C., Devare, S. G., Reddy, E. P. & Aaronson, S. A. *Science* **218**, 1131–1133 (1982).
11. Niman, H. L. & Elder, J. H. *Virology* **123**, 187–205 (1982).
12. Niman, H. L. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 4949–4953 (1983).
13. Heldin, C. H., Wasteson, A. & Westermark, B. *Expl Cell Res.* **109**, 429–437 (1977).
14. Raines, E. W. & Ross, R. *J. biol. Chem.* **257**, 5154–5160 (1982).
15. Shih, T. Y. *et al. Nature* **287**, 686–691 (1980).
16. Hunter, T. & Sefton, B. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1311–1315 (1980).
17. Kawai, S. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 6199–6203 (1980).
18. Feldman, R. A., Hanafusa, T. & Hanafusa, H. *Cell* **22**, 757–765 (1980).
19. Witte, O. N., Dasgupta, A. & Baltimore, D. *Nature* **283**, 826–831 (1980).
20. Barbacid, M., Beemon, K. & Devare, S. G. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5158–5162 (1980).
21. Neel, J. C., Ghysdael, J. & Vogt, P. K. *Virology* **109**, 223–228 (1981).
22. Feldman, R. A., Wang, L. H., Hanafusa, H. & Balduzzi, P. C. *J. Virol.* **42**, 228–236 (1982).
23. Simon, M. A., Kornby, T. B. & Bishop, J. M. *Nature* **302**, 837–839 (1983).
24. Ek, B., Westermark, B., Wasteson, A. & Heldin, C. H. *Nature* **295**, 419–420 (1982).
25. Nishimura, J., Huang, J. S. & Deuel, T. F. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4305–4307 (1982).
26. Ek, B. & Heldin, C. H. *J. biol. Chem.* **257**, 10486–10492 (1982).
27. Pike, L. K., Bowen-Pope, P. E., Ross, R. & Krebs, E. G. *J. biol. Chem.* **258**, 9383–9390 (1983).
28. Eva, A. *et al. Nature* **295**, 116–119 (1982).
29. Robbins, K. C., Antonides, H. N., Devare, S. G., Hunkapiller, M. W. & Aaronson, S. A. *Nature* **305**, 605–608 (1983).
30. Deuel, T. F. *et al. Science* **221**, 1348–1350 (1983).
31. Bittle, J. L. *et al. Nature* **298**, 30–33 (1982).
32. Green, N. *et al. Cell* **28**, 477–487 (1982).
33. Niman, H. L. & Elder, J. H. in *Monoclonal Antibodies and T Cell Products* (ed. Katz, D. H.) 23–51 (CRC Press, Boca Raton, Florida, 1982).

Primary structure homology between the product of yeast cell division control gene *CDC28* and vertebrate oncogenes

Attila T. Lörincz & Steven I. Reed

Biochemistry and Molecular Biology Section, Department of Biological Sciences, University of California, Santa Barbara, California 93106, USA

In the budding yeast, *Saccharomyces cerevisiae*, division is controlled in response to nutrient limitation¹ and in preparation for conjugation². Cells deprived of an essential nutrient or responding to mating pheromones cease division and become synchronous in the G₁ interval, apparently constrained from completing a critical event. This event has been given the operational designation of 'start'. We have isolated a large number of start mutations which confer on *S. cerevisiae* cells a conditional inability to complete start^{3,4} (Fig. 1) presumably because they define genes which must be expressed for the start event to be successfully completed. We have described the isolation on plasmids of one of the start genes, *CDC28*, by genetic complementation⁵ and initial characterization of its product^{6,7}. We now describe the DNA sequence of the gene *CDC28*.

Comparison of the amino acid sequence encoded by *CDC28* with those of several members of the vertebrate protein kinase family, including a number of viral oncogenes, reveals a significant degree of sequence homology indicating that the latter may have arisen from genes that are essential for cell proliferation and more specifically, for the normal process of cell division control.

The dideoxy method of Sanger *et al.*⁸ was used in conjunction with modified M13 phages⁹ to sequence a 1.7-kilobase (kb) segment of DNA containing the *CDC28* gene (Fig. 2) and an open reading frame of 298 codons was revealed (Fig. 3). Preliminary nuclease protection experiments¹⁰ indicate that the

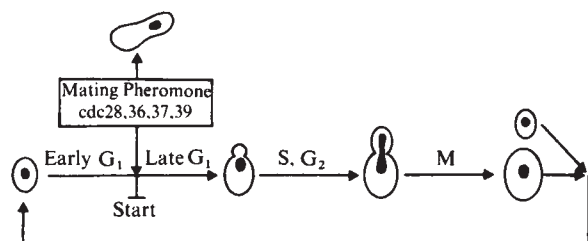


Fig. 1 The yeast cell cycle showing only the relevant details. Execution of start, the controlling event in the cell cycle, depends on both nutritional and pheromonal status. Four genes required for completion of start based on mutational analysis are *CDC28*, *CDC36*, *CDC37* and *CDC39*. Mutations at these loci produce conditional arrest at start and morphological alterations which cause mutant cells to resemble those responding to mating pheromone. Execution of start signals commitment to completion of the cell division cycle. The cell cycle intervals G_1 , S, G_2 and M are shown. The shaded body inside the cell represents the nucleus.

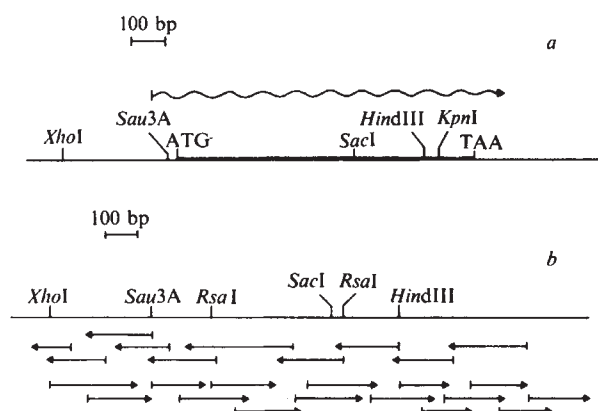


Fig. 3 The nucleotide sequence of the *CDC28* region of the yeast genome and the predicted primary structure of the *CDC28* product. The first nucleotide of the initiating codon (AUG) is given as position +1. Note that the consensus for nucleotides surrounding eukaryotic translational initiation codons is obeyed¹⁸. The predicted primary structure of the *CDC28* product has been aligned for maximal homology¹¹ with the predicted primary structure of the oncogene *v-mos* from Moloney murine sarcoma virus¹⁴ (one line down) and the primary structure of cyclicAMP-dependent protein kinase (CADPK) from bovine cardiac muscle¹³ (two lines down). This display does not necessarily constitute the maximal homology between the latter two proteins. Where gaps were necessary to obtain optimal alignment, they are indicated by leaving blank spaces (*v-mos* and CADPK sequences) or by inserting the symbol * and the number of positions comprising the gap (*CDC28* product sequence). Amino acid homologies observed between the *CDC28* product and either one of the other proteins are indicated (three lines down). Where a homology exists for all three sequences, the amino acid is followed by an asterisk. Homologies observed between only the region of homology with nucleotides long beginning at position +1 are indicated by a vertical line.

[illegible]

deviations above a mean derived by processing a large number of random sequences of identical composition and length via the same algorithm¹¹. There is no doubt, then, of the relatedness of the *CDC28* product to these diverse members of the protein kinase family, and good reason to propose that the *CDC28* product is, in fact, a protein kinase itself.

Amino acids common to *CDC28* and either of the two aligned primary structures are indicated in Fig. 3. It is interesting that even though some positions are conserved among all three sequences, many of the relationships hold in one or other of the comparisons but not both. The simplest explanation is that all three sequences are derived from the same ancestral sequence with divergence having occurred early in eukaryotic evolution. Similar matches were obtained with a number of other related viral oncogene product sequences including: Rous sarcoma virus oncogene (*src*; 21% homology), feline sarcoma virus transforming gene (*fes*; 25.5% homology), avian Fujinami sarcoma virus transforming gene (*fps*; 24.5% homology), avian Y73 sarcoma virus transforming gene (*yes*; 25% homology), mouse virus 3611 oncogene (*v-raf*; 24% homology). Most of these have been demonstrated to encode active protein kinases. For a minority, of which *v-mos* is an example, membership in the protein kinase family has been inferred from sequence data, as it has not yet been possible to demonstrate enzymatic activity *in vitro*¹⁵.

Does the sequence homology between the yeast cell cycle gene *CDC28* and viral oncogenes extend to any functional properties? Sequence comparisons between viral oncogenes and their cellular counterparts or proto-oncogenes suggest that the viral forms result from rare transduction events and that their improper expression in the infected host contributes to the transformed state¹⁶ in which normal control of cell division has been lost. An analogous situation in the yeast system is a loss of division control in the fission yeast *Schizosaccharomyces pombe* when carrying the *S. cerevisiae CDC28* gene on a plasmid¹⁷. Moreover, the *CDC28* gene can complement conditional-lethal *cdc2* mutations in *S. pombe*¹⁷ even though these two organisms are only distantly related and their mode of cell cycle control, at least superficially, dissimilar. Whether the explanation is that the *CDC28* product is unresponsive to normal *S. pombe* control signals or that the plasmid-associated abnormally high gene dosage overwhelms the regulatory capacity of the system, the behaviour of the *S. cerevisiae CDC28* gene in *S. pombe* strikes an interesting parallel with the expression of viral oncogenes in cells of higher eukaryotes and suggests that the gene products in the two yeast species have been highly conserved in the course of evolution. It is probable, then, that underlying division control mechanisms are conserved as well and that, in light of the protein kinase homology, phosphorylation of key target proteins may be involved.

We thank Dr Russell Doolittle for performing the sequence comparisons. This work was funded by NIH grant RO1 GM28005 and NSF grant PCM81-02308 to S.I.R. S.I.R. was supported in part by American Cancer Society Faculty Research Award FRA-248.

Received 24 August; accepted 24 October 1983.

1. Johnston, G. C., Pringle, J. R. & Hartwell, L. H. *Expl Cell Res.* **105**, 79–88 (1977).
2. Bucking-Throm, E., Duntze, W., Hartwell, L. H. & Manney, T. R. *Expl Cell Res.* **76**, 99–110 (1973).
3. Hartwell, L. H., Culotti, J., Pringle, J. & Reid, B. *Science* **183**, 46–51 (1974).
4. Reed, S. I. *Genetics* **95**, 561–577 (1980).
5. Nasmyth, K. A. & Reed, S. I. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2119–2123 (1980).
6. Reed, S. I., Ferguson, J. & Groppe, J. C. *Molec. cell. Biol.* **2**, 412–425 (1982).
7. Reed, S. I. *Gene* **20**, 255–265 (1982).
8. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
9. Messing, J., Crea, R. & Seeburg, P. A. *Nucleic Acids Res.* **9**, 308–321 (1981).
10. Berk, A. J. & Sharp, P. A. *Cell* **12**, 721–732 (1977).
11. Doolittle, R. F. *Science* **214**, 149–159 (1981).
12. Barker, W. C. & Dayhoff, M. O. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2836–2839 (1982).
13. Shoji, S. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 848–851 (1981).
14. Van Beveren, C. *et al. Nature* **289**, 258–262 (1981).
15. Papkoff, J., Nigg, E. A. & Hunter, T. *Cell* **33**, 161–172 (1983).
16. Duesberg, P. A. *Nature* **304**, 219–223 (1983).
17. Beach, D., Durkacz, B. & Nurse, P. *Nature* **300**, 706–709 (1982).
18. Kozak, M. *Microbiol. Rev.* **47**, 1–45 (1983).

Similarity of the *Cin1* repetitive family of *Zea mays* to eukaryotic transposable elements

Nancy S. Shepherd, Zsuzsanna Schwarz-Sommer, Jutta Blumberg vel Spalve, Manju Gupta, Udo Wienand & Heinz Saedler

Max-Planck-Institut für Züchtungsforschung, Egelspfad, D-5000 Köln 30, FRG

It has been suggested that the middle repetitive class of sequences that make up a large proportion of the eukaryotic genome have been amplified and dispersed by DNA transposition^{1,2}. Transposition is a phenomenon first postulated by Barbara McClintock on the basis of her genetic analysis of mutants in *Zea mays*³. Since then, DNA transposition has been studied genetically in various plant systems^{4–6} and is well documented on the molecular level in both prokaryotes and eukaryotes⁷. This has included the isolation of DNA inserts at various loci in several plants^{8–15}; however, the prevalence of transposition in plants is not established. We report here DNA nucleotide sequence data which show that some members of the *Cin1* middle repetitive family of maize^{16,17} have features characteristic of known transposable elements. One cloned *Cin1* repeat has a 6-base pair (bp) perfect inverted repeat sequence at its ends. The terminal five base pairs (5' TGTTG ... CAACA 3') are identical to the termini of *Drosophila copia* transposable elements¹⁸. Two other *Cin1* alleles are flanked by 5-bp direct repeats. A comparison is made with the long terminal repeat (LTR) of the *copia*-Ty1-retrovirus families of moveable genetic elements.

To define the extent of the *Cin1* element we compared the sequence of a given DNA fragment containing the insert with that of the homologous DNA fragment from which the insert is absent, using the NF1 and LC1 recombinant clones isolated from *Z. mays*. NF1 is a 5.7-kilobase (kb) clone from a Northern Flint line of maize which was found to differ from the 5.0-kb LC1 clone of a midwestern maize, line C, by the *Cin1*-001 insert (Fig. 1)¹⁹. Figure 2a shows DNA sequence analysis of the two clones in the region surrounding the *Cin1* insert. The two sequences are quite homologous (90%) except for the *Cin1*-001 insert, the left end of which is shown to be 5' TGTTGG 3' while the right end is 5' CCAACA 3', thus forming a short, 6-bp perfect inverted repeat at the ends of the element.

Further evidence that the ends of the element are assigned correctly comes from a *Teosinte guerrero* recombinant clone, TG2 (ref. 8). Annual teosintes are thought to be the progenitor of *Z. mays*; however, *T. guerrero* is not one of its closest relatives^{20,21}. The TG2 clone is homologous (but not identical) to the LC1 maize clone⁸. There is a strong correlation between the *Teosinte* sequence and that of LC1 (96%) and both show a complete absence of any portion of the sequence which we have assigned to the *Cin1* repeat (Fig. 2a).

The final approach used to define the ends of the *Cin1* element was to compare the DNA sequence surrounding the repeat with that found on other recombinant clones containing the *Cin1* repeat. Figure 2b, c shows this sequence for two such recombinant clones, LC102 and LC103; the sequences of the *Cin1* elements were clearly similar (87 and 82%, respectively; N.S.S., unpublished data), while the border regions flanking the element were completely different. The ends of the *Cin1* repeat on clone LC102 showed the identical 6-bp inverted repeat as seen for the *Cin1*-001 allele, but the LC103 clone has one base change, C→T, in this inverted repeat.

From the above analysis we conclude that the *Cin1* element terminates in a 6-bp perfect inverted repeat (5' TGTTGG ... CCAACA 3'). The presence of short, terminal inverted repeats is a characteristic of most known transposable elements; however, this particular DNA sequence is different from the

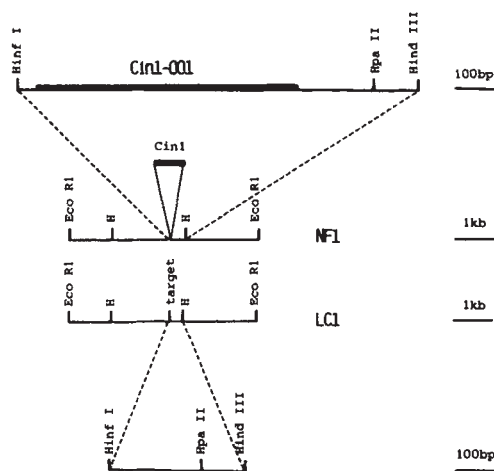


Fig. 1 The *Eco*RI and *Hind*III (H) restriction maps of the NF1 and LC1 maize recombinant clones¹⁹. The two clones were aligned to show their size similarity, except for the *Cin1*-001 sequence (thick bar) which is present on the NF1 clone.

inverted repeat sequence of other known plant DNA inserts^{8,12-14}. The terminal 5 base pairs are identical to the termini of the *copia* transposable element of *Drosophila melanogaster*¹⁸, and *Cin1* is the first plant DNA insert examined that contains the TG-CA dinucleotides present at the termini of many eukaryotic transposable elements^{7,22}.

Another characteristic of transposable elements is the duplication of target site sequence on integration⁷. This criterion is also difficult to analyse for *Cin1* because of the lack of a true target site or an immediate progenitor allele. There is a 5-bp direct repeat flanking the *Cin1* element on both the LC102 and LC103 clones (Fig. 2). The flanking sequence is not identical for the two clones (ATAAT for LC102 and GTTAG for LC103), and their presence is consistent with the notion of the *Cin1* element producing a target site duplication on integration.

Unlike the previously mentioned alleles, the *Cin1*-001 allele on the NF1 clone is not flanked by a 5-bp direct repeat. A direct repeat of 4 bp (CTCG) is present, but displaced from the *Cin1*-001 terminus on one side by two Ts. The significance of this is unknown; however, the presence of an extra base separating a direct repeat from the terminus of an insertion has been observed recently in another system²³.

The fact that the DNA sequence of the *Teosinte* clone shows greater homology with the LC1 sequence than with the NF1 sequence, does not necessarily mean that the LC1 sequence is an older form than that of NF1. *Z. mays* and annual teosintes easily interbreed and are often present in the same field^{20,21}. Therefore, it is uncertain whether or not this particular DNA fragment in the *T. guerrero* line has been exchanged with maize during growth of *Teosinte* in the wild. It is just as reasonable to assume that NF1 is the precursor of the others, and that *Cin1* had been excised from the sequence. In this case, the continued presence of the CTCG repeat plus the extra G is reminiscent of recently observed nucleotide alterations of the direct repeat on excision of the *Ds* transposable element of maize¹³. Thus, the DNA sequence of the LC102 and LC103 clones as well as that of the NF1/LC1 set of clones is consistent with the idea of *Cin1* transposition.

It has been mentioned that the TG-CA dinucleotides at the termini of *Cin1* are similar to those of other eukaryotic transposable elements (see above), including several *copia*-like elements of *Drosophila*, the Ty1 element of yeast, mouse intracisternal A-particle (IAP) genes²⁴, and the majority of retrovirus proviruses (for reviews see ref. 7). These elements contain a left and a right LTR which form a large direct repeat at the ends of the element and which also end in the dinucleotides TG-CA^{7,22,24}. Although the LTR is defined as the terminal direct repeat of a larger element, solo structures which resemble LTRs in size, hybridization or internal sequence organization, have also been found²⁵⁻²⁹.

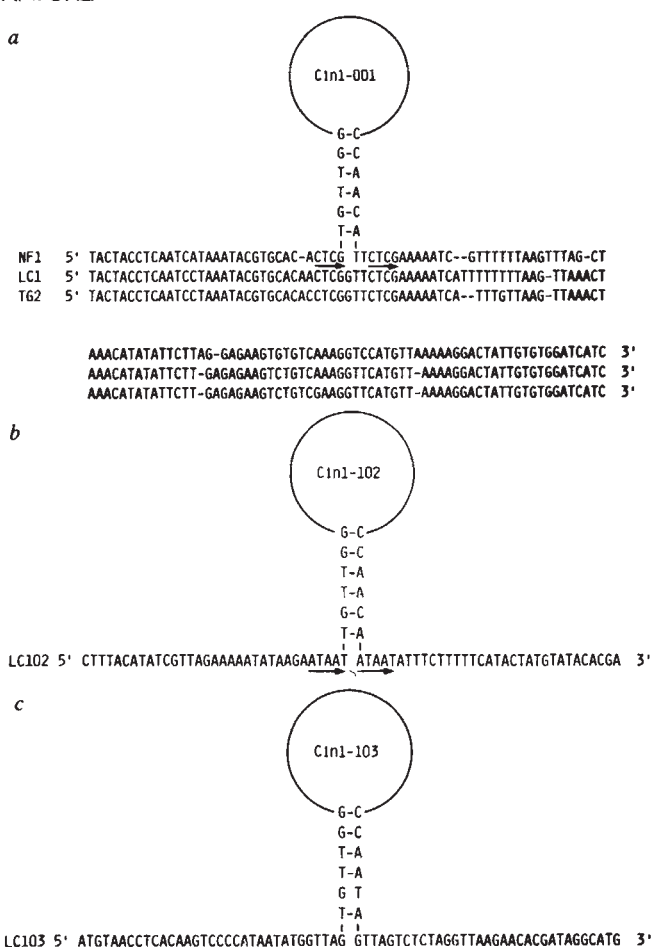


Fig. 2 DNA sequence of regions immediately flanking *Cin1* alleles present on the maize recombinant clones NF1 (a), LC102 (b), and LC103 (c)^{16,19}. *Cin1*, as defined in the text, is shown as a stem-loop structure, and direct repeats which flank the element are indicated by arrows. The corresponding homologous DNA sequences of the LC1 and TG2 recombinant clones⁸, which do not contain the *Cin1*-001 allele, are also given. The DNA sequences of NF1 (including the *Cin1*-001 sequence of Fig. 3), LC1 and LC1002 were obtained for both DNA strands using standard Maxam and Gilbert sequencing³⁹ of overlapping restriction fragments. For TG2 and the left side of LC103, only one DNA strand was sequenced, but in both cases the sequence was easily read from a 16 or 8% sequencing gel.

By analogy, the *Cin1* element is similar in size to an LTR, contains the terminal TG-CA dinucleotides, and is found dispersed in the maize genome. The internal structure of the *Cin1*-001 allele is also analogous to that of an LTR²² (Fig. 3) in that it contains signals for the initiation of RNA transcription^{30,31} followed by a poly(A) addition signal³², and signals thought to be important for the termination of RNA synthesis^{33,34}. However, it is unknown whether *Cin1* is transcribed or if it is ever found as the terminal direct repeat of a larger moveable element.

The similarities of *Cin1*-001 to the LTR of retroviruses are striking, but care should be exercised when extending this structural similarity to a mechanistic one. Retroviruses integrate into the animal genome via reverse transcription of the viral RNA genome into DNA²². The possibility of reverse transcription in plants has been the subject of many recent articles³⁵⁻³⁸ but definitive proof is yet to be provided. Whether or not these sequences are conserved in all *Cin1* alleles is unlikely, for *Cin1* has been shown to be a heterogeneous family¹⁶ and any transposition event may have occurred many years ago, thus allowing mutational drift to occur.

The findings presented here concerning the *Cin1* element of *Z. mays* suggest that it has characteristics typical of a transposable element: it is found in many copies, it has short inverted

5' TGTGGGGACCTTCTTCTTCAAGGCTCTCAAAAGCTTAATATATGTTTTCAGTATAACGAAGCTTTATAGGATGTTTCCTGTTTGGATGCAAG
 ACACCCCTGGAGAGAGAGATTTTCAGGAGTTTTCAGGATTAATATATATCAAAAGCTCATATGTTCTTGCAATCTCTACAAAGCGAAGCTCTAGTTC
 100
 ATGTGCTGATATCAAGCTCGCTTGAAGGAAGAGTGAACGAGCTGGAAGTGTGGCTGAAGGACCTTCAGCTTAGCATGATGAAGGATGTTGAAGAT
 TACAGCACTATATCTTCAAGGAGAGCTTCTTCTTCACTTCTCGAGCTTCAACAGCACTCTTGGAGTGAATGCTACTTGTACCACTTCTA
 200
 GGACGACCTCAAAACAGAAAGAGCTTCTTGTCTTGTAGCTTCTTAAATTAATTAACAATAAATCGGAGCATGAATATAATTTTCACGGCTGCG
 CTTGGCTGATTTTGTCTTCTGTAGATCAAGGAGTATCGAAGCAATTAATTAATTTATATAGCCCTGACTTATATTAAAGCTGCGCGAGCG
 300
 TCCCGGCTATAGATGAAGCAAGTACAGCTGCTTTCAGCTGCTTGAATGCTCTGATCACTCACTTCTTCACTTCTGCTCAAGCTGAAG
 AGGGCGGATTTTATCTACTTGTCTTGTGATGATGAAGTGAAGCTGATCACTTGTAGGAGGAGCTAGTGAGTGAAGTGAAGGAGCTGCTGCTC
 400
 GTATCAATGAATTAATATTTTGTGACATTTACTGATGTTTATGAATAAATGAAGATGAATATCTACTGTTATATTTGCTATTTTCACTTCTT
 CATAGTACTTATTTTATTAAGCACTGTAATGATGATCACTTATTTTATTTTCTTACATTTATAGTACCAATATACAGTAATAAATGTAAGA
 500
 CTTATGATTTTCTGCTTCTTCTTATTTATTTATTTTACCAAGCAAGGAGGATGCTTCTGATGTTTCTGCTGAAGATTTATATATCAAAAT
 GAATACATAAGTCAAGAGAGTAAATATTAATTAATTAATGTTCTGCTGCTTCCATGAGTGAAGTCAAGAGCACTTCTAATAATATAGTTT
 600
 GAGATGATGTTTTCGAAGCAAGCTATTTAATCAATCAATGTTGTTGTTTCTGATATACCACTTGAAGTAAGGAGCAACA 5'
 CTTCTACAAAGCTTCTGCTGCTGATTAATTAATGTTTGAACACAAAGAGCACTTATTTGTAAGCTTCTTCTGCTGTT

Fig. 3 DNA sequence of the *Cin1*-001 allele present on the NF1 recombinant clone. The 6-bp perfect inverted repeat at the ends of the element is marked by a thick arrow. The *Cin1* sequences that are similar to the organization of a retrovirus LTR²² are as follows. The sequence ACATT (overlined with dashes) is identical to the 'CAT' box³⁰ found in the LTR of spleen necrosis virus²². It is followed 34 bases later by a Goldberg-Hogness 'TATA' box at bases 413-419 (consensus sequence: TATA[±]A₁)³¹, a perfect poly(A) addition signal (5' AATAAA 3')³² at 467-472 (overlined), and a CA dinucleotide at position 485 (preferred site for polyadenylation, normally found 10-30 bp downstream from the poly(A) addition signal of eukaryotic genes)³². The 15-bp perfect inverted repeat (thin arrows) and the T-rich region (72% T) of the upper strand between bases 516 and 540 (brackets) are both features which may influence termination of RNA transcription^{33,34}. The longest open reading frame with an AUG near the 5' end is located on the lower strand. It begins at position 438 (ATG:*), terminates with the stop codon at position 229 (TAA:*), and possibly encodes a protein of 69 amino acids. A Goldberg-Hogness 'TATA' box is found 85 bp upstream from this open reading frame on the lower strand at bases 572-578. Other 'TATA' boxes and poly(A) addition signals are also present at various, unmarked positions. The total A+T content of *Cin1*-001 is 65%.

repeats at the end which are identical to a known transposable element, and in two cases it was found to be flanked by a 5-bp direct repeat. We conclude, therefore, that DNA transposition has played a part in the dispersion of short repetitive sequences in *Z. mays*.

We thank L. Glahn and R. Berntgen for technical assistance, and Dr Peter Czernilofsky for critical reading of the manuscript.

Received 9 August; accepted 21 November 1983.

- Britten, R. J. & Davidson, E. H. *Fedn Proc.* **35**, 2151-2157 (1976).
- d'Eustachio, P. & Ruddle, F. H. *Science* **220**, 919-924 (1983).
- McClintock, B. *Carnegie Instn Wash. Yb.* **45**, 176 (1946).
- Peterson, P. A. *Theor. appl. Genet.* **40**, 367-377 (1970).
- Fincham, J. R. S. & Sastry, G. R. K. *A. Rev. Genet.* **8**, 15-50 (1974).
- Fincham, J. R. S. *Genetics Suppl.* **73**, 195-205 (1973).
- Shapiro, J. A. (ed.) *Mobile Genetic Elements* (Academic, New York, 1983).
- Saeder, H. et al. in *Genetic Rearrangement, 5th John Innes Symp.* (eds Chater, K. F. et al.) 107-115 (Croom Helm, London, 1983).
- Geiser, M. et al. *EMBO J.* **1**, 1455-1460 (1982).
- Burr, B. & Burr, F. A. *Cell* **29**, 977-986 (1982).
- Strommer, J. N. et al. *Nature* **300**, 542-544 (1982).
- Goldberg, R. B., Hoschek, G. & Vodkin, L. O. *Cell* **33**, 465-475 (1983).
- Peacock, W. J. et al. in *15th Miami Winter Symp.* (1983).
- Döring, H. P., Tillman, E. & Starlinger, P. *Nature* **307**, 127-130 (1984).
- Fedoroff, N., Wessler, S. & Shure, M. *Cell* **35**, 235-242 (1983).
- Gupta, M., Bertram, I., Shepherd, N. S. & Saeder, H. *Molec. gen. Genet.* **192**, 373-377 (1983).
- Gupta, M., Shepherd, N. S., Bertram, I. & Saeder, H. *EMBO J.* (in the press).
- Levis, R., Dunsmuir, P. & Rubin, G. M. *Cell* **21**, 581-588 (1980).
- Shepherd, N. S. et al. *Molec. gen. Genet.* **188**, 266-271 (1982).
- Illis, H. H. & Doebley, J. F. *Am. J. Bot.* **67**, 994-1004 (1980).
- Wilkes, H. G. *Science* **177**, 1071-1077 (1972).
- Weiss, R., Teich, N., Varmus, H. & Coffin, J. *RNA Tumor Viruses* (Cold Spring Harbor Laboratory, New York, 1982).
- Rechavi, G., Givol, D. & Canaan, E. *Nature* **300**, 607-611 (1982).
- Kuff, E. L. et al. *Nature* **302**, 547-548 (1983).
- Cole, M. D., Ono, M. & Huang, R. C. C. *J. Virol.* **38**, 680-687 (1981).
- Cameron, J. R., Lok, E. Y. & Davis, R. W. *Cell* **16**, 739-751 (1979).
- Hughes, S. H., Toyoshima, K., Bishop, J. M. & Varmus, H. E. *Virology* **108**, 189-207 (1981).
- Brown, S. D. M. *Gene* **23**, 95-97 (1983).
- Streeck, R. E. *Nature* **298**, 767-769 (1982).
- Elstratidis, A. et al. *Cell* **21**, 653-668 (1980).
- Corden, J. et al. *Science* **209**, 1406-1414 (1980).
- Proudfoot, N. J. & Brownlee, G. G. *Nature* **252**, 359-362 (1974).
- Rosenberg, M. & Court, D. A. *Rev. Genet.* **13**, 319-353 (1979).
- Lai, E. C. et al. *Cell* **18**, 829-842 (1979).
- Varmus, H. *Nature* **304**, 116-117 (1983).
- Hull, R. & Covey, S. N. *Trends Biol. Sci.* **8**, 119-121 (1983).
- Pfeiffer, P. & Hohn, T. *Cell* **33**, 781-789 (1983).
- Toh, H., Hayashida, H. & Miyata, T. *Nature* **305**, 827-829 (1983).
- Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499-560 (1980).

A large increase in enzyme-substrate affinity by protein engineering

Anthony J. Wilkinson*, Alan R. Fersht*, D. M. Blow†, Paul Carter‡ & Greg Winter‡

* Department of Chemistry and † Department of Physics, Imperial College of Science and Technology, London SW7 2AY, UK

‡ MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

A single point mutation has been engineered in the tyrosyl-tRNA synthetase that improves its affinity (K_M) for its substrate ATP by a factor of 100. In the crystal structure of the tyrosyl tRNA synthetase (of *Bacillus stearothermophilus*), the side-chain hydroxyl of Thr 51 appears to make a weak hydrogen bond with the AMP moiety of the substrate intermediate, tyrosyl adenylate. In the absence of substrate, however, the hydroxyl group should make a strong hydrogen bond with water which would favour dissociation of the enzyme-substrate complex. We have used oligodeoxynucleotide-directed mutagenesis to construct two point mutants at this site: one to remove the hydroxyl group (Thr 51 → Ala 51) and the other, in addition, to distort the local polypeptide backbone (Thr 51 → Pro 51). We report here that both mutants have increased activity (k_{cat}/K_M for ATP) but one mutant (Pro 51) shows a massive 25-fold increase due mainly to a lowered K_M for ATP. This demonstrates dramatically the potential of *in vitro* mutagenesis for improving the affinity of an enzyme for its substrate.

The activities of such bacterial enzymes as β -lactamase¹, 'evolved β -galactosidase'^{2,3}, amidases^{4,5} and ribitol dehydrogenase⁶ have been successfully directed towards novel substrates *in vivo* by selecting spontaneous mutations of the enzyme. We are exploring an alternative approach in which we inspect a crystallographic model of the enzyme and attempt to predict the effect of changing a side chain in the active site. We then construct the required mutant by site-directed mutagenesis of the gene, express the enzyme in *Escherichia coli* and measure catalytic activity^{7,8}.

The tyrosyl-tRNA synthetase (TyrTS) of *B. stearothermophilus* catalyses the aminoacylation of tRNA^{Tyr} in a two-step mechanism in which tyrosine is activated to form enzyme-bound tyrosyl adenylate and then transferred to tRNA^{Tyr}. The location of tyrosyl adenylate in the crystal structure indicates that hydrogen bonds are made to the ribose moiety from the side chains of Cys 35, Thr 51 and His 48 (refs 9-11). The Cys 35 interaction has been confirmed by constructing Gly 35 and Ser 35 mutants: these have lowered affinities for ATP resulting either from the complete removal of the hydrogen bond in TyrTS(Gly 35) or from placing the -OH group in TyrTS(Ser 35) slightly too far from the hydrogen bond acceptor on the ATP^{7,8}. Interestingly, TyrTS(Ser 35) has a lower affinity for ATP (calculated from k_{cat}/K_M for aminoacylation) than does TyrTS(Gly 35), despite the possibility of a weak hydrogen bond. This probably stems from Ser 35 exchanging a good hydrogen bond with water in the free enzyme for a poor one with ATP on forming the enzyme-substrate complex⁸. As the hydroxyl group of Thr 51 makes a long hydrogen-bonding contact to the ring oxygen (O-1) of the ribose, we predicted that removal of the weak hydrogen bond by constructing an Ala 51 mutant might improve the affinity of the enzyme for ATP.

The Ala 51 mutant was constructed as described previously^{7,12} (using a synthetic oligodeoxynucleotide, 5' AAATGGCGGCCAAGTG 3') and characterized by dideoxy sequencing¹³. Not only was the Ala 51 lesion confirmed, but also the entire TyrTS gene was sequenced to prove that no other mutations had been introduced inadvertently. For this purpose, a family of five synthetic oligodeoxynucleotide primers staged throughout the TyrTS gene was used (Fig. 1A), allowing the complete sequence to be determined on a single salt gradient

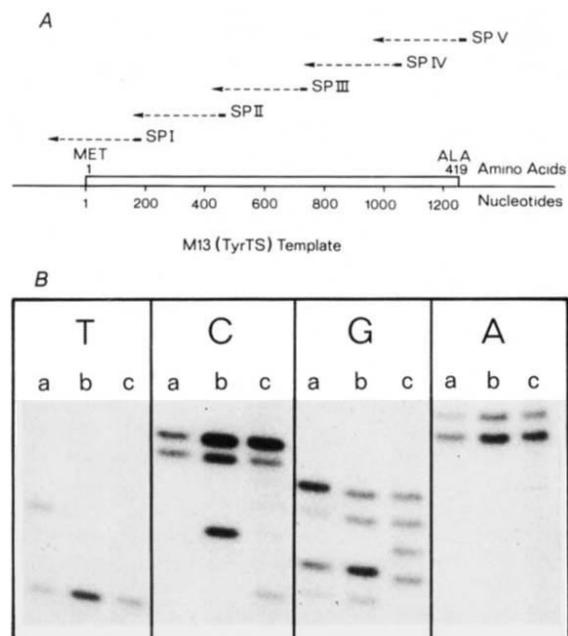


Fig. 1 Sequencing of mutant TyrTS gene. **A**, Oligodeoxynucleotide primers: SPI, 5' CCCGCTGCTGGAAGCG 3'; SPII, 5' TGAACCGACTCTTT 3'; SPIII, 5' CTCGTACGGCGACGTTT 3'; SPIV, 5' AAGCGGAACGTCGCCTC 3'; SPV, 5' TTACTATGCCAGCGC 3', complementary to the sequence of the TyrTS gene¹⁵ cloned in M13, were synthesized using phosphotriester chemistry on a polydimethylacrylamide kieselguhr support¹⁶ and purified by ion-exchange HPLC. Templates from the mutants M13TyrTS(Ala 51) and M13TyrTS(Pro 51) were prepared¹⁷ and the genes sequenced by primer extension and the dideoxy chain termination method. As >300 nucleotides from the primer can be read from gradient gels using [α -³²S]thio-dATP¹⁴, the complete sequence of each gene may be determined in a single gel. **B**, Sequences of M13TyrTS(Thr 51) (tracks a), M13TyrTS(Ala 51) (tracks b) and M13TyrTS(Pro 51) (tracks c) as determined using the primer SPI. Detection of sequence changes was facilitated by running the T, C, G or A tracks of the three clones in adjacent slots.

gel¹⁴. By running together the corresponding T, C, G and A tracks of mutant and wild-type clones, sequence changes can be readily detected (Fig. 1B).

Mutation of Thr 51 to Ala 51 causes a small change in k_{cat} for both exchange (Table 1) and aminoacylation (Table 2), but decreases K_M for ATP by a factor of two for both reactions. The specificity constant, k_{cat}/K_M , is accordingly increased two-fold, indicating that the enzyme-substrate interaction energy is increased by 0.38 kcal mol⁻¹ in the Ala 51 mutant. This is consistent with our prediction that a small increase in enzyme affinity might be engineered by deleting a poor hydrogen bond with the substrate.

Residues 47–61 all have ϕ , ψ angles in the range characteristic of an α -helix (-57° , -48°)¹¹, but the helix is slightly irregular at the amino-terminus, and the hydrogen bond between Gly 47 (CO) and Thr 51 (NH) may not form. In the homologous *E. coli* enzyme where residue 51 is proline¹⁵, this hydrogen bond could not exist as the nitrogen is locked into the pyrrolidine ring. This must cause a small distortion in the polypeptide backbone around residues 47–51. To examine the effect on the affinity of ATP, we constructed a Pro 51 mutant of the *B. stearothermophilus* enzyme (as above, using the primer 5' CAAAATCGGGGCCAAGT 3'). The mutation leads to large kinetic changes. In the activation reaction, there is a 15-fold lowering of K_M for ATP and a near doubling of the value of k_{cat} . These combine to give a 25-fold increase in the specificity constant (≈ 1.9 kcal mol⁻¹ of activation energy). In the charging reaction, K_M for ATP is reduced 130-fold for TyrTS(Pro 51), but k_{cat} is also reduced by a factor of 2.6 so that the specificity constant is 50 times higher than that for the wild type (≈ 2.4 kcal

Table 1 Pyrophosphate exchange activity of tyrosyl-tRNA synthetases

Enzyme	k_{cat} (s ⁻¹)	K_M (ATP) (mM)	k_{cat}/K_M (s ⁻¹ M ⁻¹)
TyrTS	7.6	0.9	8,400
TyrTS(Ala 51)	8.6	0.54	15,900
TyrTS(Pro 51)	12.0	0.058	208,000

Synthetase exchange activity was measured at 25 °C, pH 7.8, in 144 mM Tris-HCl, 10 mM MgCl₂, 0.1 mM phenylmethane sulphonyl chloride, 10 mM 2-mercaptoethanol, 2 mM pyrophosphate, 50 μ M tyrosine, and 100–250 nM enzyme (assayed by active site titration).

Table 2 Aminoacylation activity of tyrosyl-tRNA synthetases

Enzyme	k_{cat} (s ⁻¹)	K_M (ATP) (mM)	k_{cat}/K_M (s ⁻¹ M ⁻¹)
TyrTS	4.7	2.5	1,860
TyrTS(Ala 51)	4.0	1.2	3,200
TyrTS(Pro 51)	1.8	0.019	95,800

The same buffer was used as for Table 1, plus 100 μ M [¹⁴C]-Tyr, 20 μ M tRNA^{Tyr} and 25–50 nM enzyme.

mol⁻¹ of activation energy). The detailed stereochemical explanation of this mutation must await further work: for example, it is conceivable that the distortion of the polypeptide backbone introduced by Pro 51 may alter the position of His 48, and increase its interaction with the substrate. However, one factor in the improved affinity must be the removal of the poor hydrogen bond of Thr 51 with ATP.

Would the mutation to TyrTS(Pro 51) confer a selective advantage on the cell, given its improved K_M with ATP? In addition to any change in other properties of the protein that have as yet not been investigated, such as thermostability, there are clear kinetic disadvantages of TyrTS(Pro 51) *in vivo* because of the high cellular ATP concentration. The intracellular concentration of ATP (2–3 mM) is so high that the wild-type and mutant enzymes are nearly saturated with ATP, thus the reaction rate is controlled by k_{cat} and not k_{cat}/K_M (ref. 8). This confers a slight disadvantage on TyrTS(Ala 51) and a greater disadvantage on TyrTS(Pro 51) in the overall charging reaction (Table 2). However, *in vitro*, TyrTS(Pro 51) is far more active at lower ATP concentration (<2 mM). Also, in the tyrosine activation reaction, TyrTS(Pro 51) is far more active at all concentrations of ATP. (Selective pressure is exerted on the overall aminoacylation reaction rather than on the partial reaction of activation.)

Irrespective of the consequences *in vivo*, the dramatic improvement of K_M for ATP *in vitro* has significant implications for industrial enzymology by demonstrating that it is possible to engineer large improvements in enzyme activity by site-directed mutagenesis.

Received 16 September; accepted 4 November 1983.

- Hall, A. & Knowles, J. R. *Nature* **264**, 803–804 (1976).
- Hall, B. G. *Genetics* **89**, 453–465 (1978).
- Hall, B. G. & Zuzel, T. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3529–3533 (1980).
- Paterson, A. & Clarke, P. H. *J. gen. Microbiol.* **114**, 75–85 (1979).
- Turberville, C. & Clarke, P. H. *FEMS Microbiol. Lett.* **10**, 87–90 (1981).
- Wu, T. T., Lin, E. C. C. & Tanaka, S. *J. Bact.* **96**, 447–456 (1968).
- Winter, G., Fersht, A. R., Wilkinson, A. J., Zoller, M. & Smith, M. *Nature* **299**, 756–758 (1982).
- Wilkinson, A. J., Fersht, A. R., Blow, D. M. & Winter, G. *Biochemistry* **22**, 3581–3586 (1983).
- Monteilhet, C. & Blow, D. M. *J. molec. Biol.* **122**, 407–417 (1978).
- Rubin, J. & Blow, D. M. *J. molec. Biol.* **145**, 489–500 (1981).
- Bhat, T. N., Blow, D. M., Brick, P. & Nyborg, J. *J. molec. Biol.* **158**, 699–709 (1982).
- Zoller, M. & Smith, M. *Nucleic Acids Res.* **10**, 6487–6500 (1982).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
- Biggin, M. D., Gibson, T. J. & Hong, G. F. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3963–3965 (1983).
- Winter, G., Koch, G. L. E., Hartley, B. S. & Barker, D. G. *Eur. J. Biochem.* **132**, 383–387 (1983).
- Gait, M. J., Matthes, H. W. D., Singh, M., Sproat, B. S. & Titmas, R. C. *Nucleic Acids Res.* **10**, 6243–6254 (1982).
- Winter, G. & Fields, S. *Nucleic Acids Res.* **8**, 1965–1974 (1980).

Mathematical pragmatism

Stephen Toulmin

The Nature of Mathematical Knowledge. By Philip Kitcher.
Oxford University Press: 1983. Pp.287. £15, \$25.

IT WAS more than time for a shift of focus in the philosophy of mathematics. With very few exceptions (Henri Poincaré or Hugo Dingler, for example) twentieth-century writers in this field have worked under the shadow of Gottlob Frege; and Frege's classic work, *The Foundations of Arithmetic* (1884), apparently demolished for good the attempt to give mathematics a basis in our empirical experience. In particular, Frege's denunciations of "the genetic fallacy", and of "psychologism", were part of a larger, Platonizing attack on the whole subject, and they have been understood to rule out any hope of introducing historical and mental factors into the analysis of mathematical concepts and propositions.

Since Frege, there have been few truly radical changes in the subject. For all of his originality, Ludwig Wittgenstein's posthumous *Remarks on the Foundations of Mathematics* (1967) never wholly emerged from under Frege's shadow. And although Imre Lakatos's early *Proofs and Refutations* (1963–1964) placed some emphasis on the historical mutability of basic mathematical notions like "validity", "rigour" and "truth", his alliance with Karl Popper led him, later on, to downplay those insights. Instead, he moved toward a more timeless view of the subject matter of mathematics, as part of a distinct "Third World", set apart from all changing physical and mental things.

Although Frege was a contemporary of Ernst Mach, who showed how far the basic concepts of physical theory are the products of human effort over time — so that, as many philosophers of science today would agree, philosophical accounts of scientific concepts need to be *historisch-kritisch dargestellt* — the corresponding insight in the philosophy of mathematics has hitherto made little headway. Nowadays, philosophers rarely focus on the empirical roots of mathematical concepts and procedures, because they have set themselves an oversimplified choice: either, they assume, mathematical truths must be based on empirical experience in the present — the view attributed to John Stuart Mill, which Frege demolished so devastatingly — or, in reaction, they deny to mathematics any empirical basis beyond the self evidence of its axioms and, like Frege himself, treat its claims as transhistorical.

Yet a powerful argument can be presented for thinking that these are not the only options. Must our understanding of the procedures and concepts of mathe-

matics, and even their "meaning", rest solely on observations and experiences that can be pointed to *in the present*? Or do they not depend, equally, on those *past* experiences that led earlier mathematicians to adopt the systems and formalisms they actually did, rather than other systems or formalisms, which were equally consistent and conceivable? This alternative philosophical approach need not, of course, be taken to imply that mathematical statements function as simple reports of past experience. Rather, it asks whether we can fully grasp the intellectual force and scope of mathematical theories and theorems, if we look only at their formal features, in isolation both from the historical contexts in which they took the forms they did, and from the psychological contexts in which children can master them in the present. This richer alternative is one that presses itself on us the more urgently, at a time when philosophers of science have moved so far towards an alliance with historians of scientific thought, and when the developmental studies of L.S. Vygotsky and Jean Piaget have redirected our attention towards the learning of mathematical concepts and procedures.

Philip Kitcher's book, *The Nature of Mathematical Knowledge*, is a pioneer attempt to give a thorough philosophical account of mathematics from this third, and novel, point of view. As is the case with many such enterprises, the book opens up more issues than it can cover to general satisfaction. Still, it has many striking features. It opens with a dozen pages of introduction, which provide a clear and courageous statement of the central programme of the new approach, and this deserves a critical reading by all philosophers interested in the subject, quite aside from the rest of the text. (For instance, the introduction hints at many connections that the author does not pursue further: e.g. at parallels between his view of mathematics and J.J. Gibson's "ecological realism" as a view of sense perception.) Among other good things, subsequent chapters explore helpfully the relations between the "historical-critical" account of change in natural science that has developed in recent years, and the novel picture of historical change in mathematical concepts and procedures that will have to be built up, if the new approach establishes itself as a new research programme for the philosophy of mathematics.

Professor Kitcher's argument falls into three parts. Four chapters clear away philosophical undergrowth, and prepare the

ground for the constructive work which occupies the author in Chapters 5 to 9, while a long final chapter is given over to a careful study of the development of "Analysis" as a formal branch of mathematics. (I first wrote "a branch of *pure* mathematics", but struck this out, for Professor Kitcher's approach undermines the traditional contrast between pure and applied mathematics. From his point of view, the only distinctions are ones of degree: based on the comparative distances of different mathematical concepts and theorems from the practical contexts and experiences out of which they initially crystallized.)

Recently, debate within the philosophy of mathematics has become rather specialized and inbred: as a result, those not up in the subject may find the opening part of the book obscure. But such readers are urged to turn to Chapter 5, where much in the constructive exposition of the new position is presented directly and clearly, without reference to current professional controversies. In particular, Professor Kitcher's discussion of the historical character of conceptual change in natural science and mathematics should be read and discussed by many people — working mathematicians and general philosophers, as well as others — who have no particular commitment to one or another of the current rival positions in the philosophy of mathematics.

If this book is in the end somewhat tantalizing, it is for the best of reasons: viz., that the author throws out many more novel ideas and suggestions than he has the space to follow up. Still, at a few points, this gap between promise and achievements affects his argument. To give just one example, his reading of psychology is limited to certain "individualist" traditions: apart from his passing reference to J.J. Gibson, the other main references are (for example) to J.A. Fodor and Noam Chomsky. Hence, for instance, he treats the *warranting* of a belief as a "mental process" in the head of an individual thinker, and ignores the alternative view that treats *warrants* rather as elements within collective procedures of argument that are developed and applied collectively, or "in the public domain".

Minor shortcomings apart, however, this is a valuable and important book. After the careful and thorough criticism of "creationism" in his earlier book, *Abusing Science*, it is a pleasure to see Philip Kitcher making such an original contribution to the philosophy of mathematics, and to general philosophy also. All told, it marks him as having one of the freshest and most versatile minds at work in Anglo-American philosophy today. □

Stephen Toulmin is a member of the Committee on Social Thought at the University of Chicago. His most recent book is The Return to Cosmology: Postmodern Science and the Theology of Nature (University of California Press, 1983).

All the way through mammalian skin

Sam Shuster

Mammal Skin.

By V.E. Sokolov.

University of California Press: 1983.

Pp.695. £61.50, \$72.50.

AFTER many dark years, the academic wilderness of the outer cover has been blown and biomedical scientists are coming out of the warmth of internal prying to the rewarding uncertainties of the cold surface. The spate of cutaneous publishing is part of the celebration and in that sense it may be churlish to ask for sober assessment rather than whoopee.

As might have been expected from a field in which the drive has mostly been pragmatic (disease, fur and pelt production), some topics have been dealt with rather more than others. As might also be expected of an organ whose great complexity is the unique issue of a commune rather than a marriage of other tissues, expertise can easily be maintained in one aspect side by side with the deepest oblivion of neighbourhood experts working on a different topic not a μ or two away. From this it follows that most large skin texts are multi-author jobs and they, and we, all suffer from it. The thrust of this over-long

welcome to this singular book must by now be apparent: we have for the first time in the English language a notably unique and personal account of the comparative structure of the mammalian skin.

There is a short general introduction on structure followed by a systematic description covering the entire immense range of mammalian skin, the book ending with a section on adaptive changes. The whole is so exhaustive that you can follow your pet themes right through their evolutionary vicissitudes, and, if you've the stamina left, crack your equally pet theories against reality: you will not travel far without finding some extraordinary and interesting new facts which, at very least, will give you a clutch of new experiments for students. This then is an old-fashioned work of scholarship the like of which is all too rarely seen.

Nonetheless there are some serious

defects. The thinking, information and methodology are essentially of the 1960s and earlier, although the original 1974 Russian edition is said to have been updated (in compensation it is salutary to be reminded how well simple methods work — for instance the evidence for the thermo-regulatory function of the fins of certain aquatic mammals); some of the data is thin; and there are some surprising omissions — coat colour and pattern is an opportunity lost. Yet despite the many flaws, including a rudimentary index and indescribably bad photomicrographs, this is an outstanding contribution to the literature. But I do hope that next time round the book gets a thorough revision to bring it into the 1980s; it certainly deserves it. □

Sam Shuster is Professor of Dermatology at the University of Newcastle upon Tyne.

Superficial science

J.E. Inglesfield

Surface Physics, 2nd Edn.

By M. Prutton.

Clarendon: 1983. Pp.138.

Hbk £12.50; pbk £6.95.

EVEN before starting to study the chemisorption of carbon monoxide on nickel, a fledgling surface scientist (unless he is a theoretician) will have to be familiar with vacuum techniques, Auger spectroscopy and the theory of low-energy electron diffraction. Professor Prutton's *Surface Physics* is an ideal guide for such a student in his first few terms, introducing him to the techniques used for surface analysis and structure determination, and the main concepts in surface electronic structure, thermal motion of surface atoms and adsorption. Detail is eschewed, but the account is commendably clear and there is an up-to-date list of references.

One of the remarkable features of surface science is the variety of techniques which are routinely used to unravel surface structure. This, quite justifiably, is the topic of the longest chapter in the book, and brief but lucid accounts are given of low-energy electron diffraction (LEED), reflection high-energy electron diffraction (RHEED), field ion microscopy (FIM), ion scattering, photoelectron diffraction and surface-extended X-ray absorption fine structure (SEXAFS). These techniques are based on a whole range of physics, from solving the electronic Schrödinger equation in LEED and SEXAFS to following the classical trajectories of Ar^+ ions with kilovolt energies in ion scattering — the physics of the techniques is briefly but clearly explained. We can expect more techniques (and more confounded acronyms) in the third edition — surely on the way are grazing incidence X-ray

diffraction, low-energy atom diffraction, and most impressive of all, the scanning tunnelling microscope.

The drawback of a book such as this is that it is inevitably bitty. This makes me less sure about its use by undergraduates. An excellent final-year undergraduate course can certainly be based on surface science, as it provides so many examples of physical principles — particularly quantum mechanics — in action. For example, the electron beam in a LEED experiment will be completely reflected by the surface when its energy coincides with a forbidden energy gap in the bulk band structure, because at such an energy electrons cannot propagate inside the solid. The concept of energy gaps leads to surface states — wavefunctions localized at the surface because their energies lie in a bulk gap below the vacuum zero of energy — of which the dangling bonds on a semiconductor surface provide one example. In the space available Professor Prutton can only touch on topics such as surface states, and the unifying role of surface electronic structure has to be neglected. Undergraduates will find the book useful when they want to find out how surface states have been found experimentally.

Anyone beginning a career in a surface science lab will find this little book invaluable. Theoreticians can survive a bit longer than experimentalists without knowing what an ultra-high vacuum is, or what a LEED apparatus looks like, but they cannot get by indefinitely and they too ought to read this book when they begin research. One of the real delights of surface science is that it is still possible to understand most of what's going on, and this book provides an excellent piece of elementary education in the subject. □

J.E. Inglesfield is Head of the Theory and Computational Science Division at the Science and Engineering Research Council's Daresbury Laboratory, Warrington.



AMERICAN NUCLEAR SOCIETY'S PUBLICATIONS

Complete Nuclear Publications and Services Worldwide.

Inquire about specific interests.

For Free Publications and Services Catalog, contact:

American Nuclear Society
555 N. Kensington Avenue
La Grange Park, IL 60525
U.S.A.

Meeting of lizards and ecologists

Mark Williamson

Lizard Ecology: Studies of a Model Organism.

Edited by Raymond B. Huey, Eric R. Pianka and Thomas W. Schoener.
Harvard University Press: 1983. Pp.501.
\$35, £29.75.

FRASER Darling used to advise that the way to start writing is to put a semi-colon in the middle of a blank page. A similar technique has been used in this survey. Most of the chapters started as papers given at a symposium in December 1980, and these have been topped and tailed with introductions to the three main sections and to the book as a whole, and a final overview. The intention is to cover not all but several aspects of lizard ecology, evaluating progress in the past fifteen years.

Much of modern ecological theory stems from the study of birds, particularly from the work of David Lack and Robert MacArthur. The title, and some of the comments, suggests that lizards could be equivalent to birds as model organisms for understanding ecology. They are certainly different, less diverse but far from uniform. The major difference is that lizards are low-energy animals, with the advantages of being easy to observe and capture, hardy in captivity, and many species, particularly in the American west, are abundant. But although the low-energy theme recurs through the papers it is a muted refrain; most authors expound their subject for its own sake, and only by inference compare their subject with the high-energy birds and mammals.

The three sections cover the main aspects of modern ecology. The first quarter of the book is on physiological ecology, an overview and four weakly co-ordinated chapters in different styles. Of the three chapters concerned with energy, one is a long and experimentally sophisticated field study of one species (*Uta stansburiana*), one a short survey of locomotion and metabolism in the laboratory, the third a not very successful attempt to understand distributional limits from the effect of temperature on activity. The other chapter in this section is on the physiological consequences of lizard malaria; this disease causes roughly a 20% decrease in physiological performance, but it may not affect population density. The rest of the book is similar in style, a mixture of detailed studies, broad surveys and loose ends.

The second section, another quarter of the book, is on behavioural ecology and is much concerned with the slippery topic of home range, in *Uta* again, and in relation to sex dimorphism. A short chapter by Regal on adaptive zone and behaviour does address the question of how lizards differ

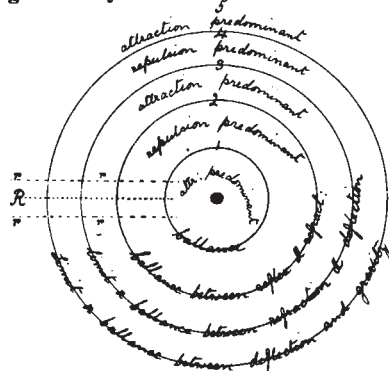
from birds and mammals, the physiological and anatomical limitations of ecological performance, and the diversity of life styles in lizards. Other chapters are less successful.

That leaves half the book for population and community ecology, which is certainly not unreasonable in relation to the excitement of work in those fields in the past couple of decades. The section starts with a sound and helpful overview by Thomas Schoener, and ends with a personal, and more amusing, conclusion by the ornithologist Peter Grant. In between, half the space is taken by *Anolis* in the West Indies. On the Greater Antilles this genus has had a most interesting adaptive radiation, and Ernest Williams extends his previous study on Puerto Rico to encompass Hispaniola and Jamaica, around fifty species in all. In contrast, on the Lesser Antilles there are never more than two species on any one island, and Roughgarden, Heckel and Fuentes consider whether there has been co-evolution of competing populations. Their empirical evidence is reasonably convincing, their mathematical modelling less so. But certainly studies on *Anolis* are important for all ecologists, evolutionists and biogeographers. Chapters by Dunham, by Huey and Pianka, and by Case are also concerned with the measurement of competition, and whether niche overlap is a quantifiable concept. This is an active and controversial field, and the papers here make a useful contribution.

Overall this volume is as patchy as most reports on symposia; the attempts at cohesion are rather flimsy. For instance, the references are grouped at the end, but arranged by chapter and not unified. Lizards have much to tell ecologists, but for the moment those outside the field will have to make their own assessment of the extent to which knowledge of lizard ecology bears on their interests. □

Mark Williamson is Professor of Biology at the University of York.

Light history



William Herschel's diagram of 1780, showing his conception of the force field surrounding a particle of matter. The illustration is taken from *Optics after Newton*, by Geoffrey Cantor, recently published by Manchester University Press. Price is £20.

More or less belief in reproduction

John Habgood

The Biology of Religion.

By V. Reynolds and R.E.S. Tanner.

Longman: 1983. Pp.332. £14.95, \$30.

THOSE who enjoy out-of-the-way information will find plenty to entertain them in this fascinating book. Where else could one lay one's hand on comparative figures for obesity among Trappist and Benedictine monks? Or on the Tibetan custom of drinking tea out of a cauldron which has just held a corpse? Or find a discussion of the relationship between wet-nursing, infant mortality and Christian beliefs about sexuality in seventeenth-century France? There are paragraphs here on everything from Islamic divorce to religious massacres. The hugely varied religious beliefs and practices of mankind have been combed for anything which might be thought to have a bearing on the health, survival and reproductive potential of those who profess them. The result is a book which is a delight to read, or to browse in, and which assembles an uncommonly comprehensive range of scientific studies on religious practice across the world.

The bulk of the material is arranged in terms of the human life-cycle, and the religious rituals and moral attitudes associated with each stage in human development in different religious cultures are assessed in terms of their reproductive consequences. There is a somewhat shorter section which tackles specifically the relationship between religion and disease. The results are not always what the protagonists of different religions might wish to believe. There is an uncomfortable suggestion, for instance, that the Flagellants, whose response to the Black Death in the fourteenth century was to go around the countryside flogging themselves to demonstrate their penitence, were actually instrumental in spreading infection.

The layout of the book, with its wide margins and numerous line-drawings, might give the impression that it is designed more for the coffee table than as a serious scientific study. But although the authors hope that general readers will be interested, as surely they will, their main purpose is to expound and test a hypothesis about cultural evolution, using religion as a convenient cultural indicator. It is on this level, therefore, and not as an entertaining compendium, that the work must ultimately be judged.

In the authors' own words, the hypothesis is that the more unpredictable the total environment is perceived to be, the more people will

want to create and support ideas favouring high levels of reproduction. They will accept the loss of some of their children as inevitable. Conversely, the less unpredictable the environment is perceived to be, the more people will develop ideas favouring low levels of reproduction.

The hypothesis is tested, in the particular case of religion, by classifying different religious strategies as reproductively negative or positive, and then looking for correlation between broad religious traditions and such indicators of environmental stability as energy consumption and gross national product. Not surprisingly, Protestant Christians turn out to be wealthy energy-consumers with inhibited attitudes towards sex, whereas Hindus are just the opposite.

Such a bald summary does not do justice to the considerable subtlety often exercised in the interpretation of religious ideas — the perceptive discussion of celibacy is a

case in point — nor to the sensitive exploration of ways in which cultural and physical evolution might in practice interrelate. I welcomed also the frank recognition that there is vastly more to religion than the biologically adaptive dimension on which the authors concentrate. But even within its own terms the main hypothesis seems to me to be thin and unconvincing. I was surprised, for example, to find no mention of the widely canvassed view that Protestant Christianity, far from being a response to an environment perceived as stable, has been a major factor in helping to bring that environment under control. In other words, it is not only the relationships between causes and effects which are uncertain in this immensely complicated field; there is also the question of which is which. □

John Habgood is Archbishop of York. Earlier in his career he was a physiologist and pharmacologist at the University of Cambridge.

Hoyle: comparative excellence

P.N.R. Usherwood

Muscles and their Neural Control.

By Graham Hoyle.
Wiley: 1983. Pp. 688.
£61.30, \$85.85.

GRAHAM Hoyle's monograph, *Comparative Physiology of the Nervous Control of Muscular Contraction*, was required reading in the late 1950s and early 1960s when comparative neurobiology was still a fledgling undergraduate subject. At that time "according to Hoyle", took on an entirely new meaning. After almost 30 years, the revised version, *Muscles and their Neural Control*, has appeared. It is well written, beautifully illustrated and conveys a most comprehensive, if at times birds-eye view of muscle structure and function across the animal kingdom. It is a remarkable achievement, but hardly a surprising one as the author stands out amongst neurobiologists for the breadth of his research and the length of his bench career. Representing as it does a distillation of his life's work, it commemorates an era which is now sadly slipping away from us.

The book is intended for undergraduates and is at its best when describing those muscle preparations which formed the stepping stones of the author's career — there are numerous references to published and unpublished work of Hoyle and his students — but the style falters when less familiar territory is encountered. His description of muscle fine structure is closely linked to cogent functional correlates and, together with a detailed phylogenetic survey of muscle innervation, this provides a successful approach for those

seeking a comparative perspective of muscle.

This account of contemporary knowledge and ideas in nerve-muscle physiology is laced with words of caution from a scientist who clearly resents the oft pejorative interpretation of the term comparative. He is also unhappy about current views on the molecular mechanisms of contraction which have emanated mainly from studies of vertebrate muscle. Past challenges to the cross-bridge hypothesis fashioned from his discovery, or rediscovery, of ultrathin filaments in some muscles have met "a very sceptical, unreceptive audience bent solely upon tracking down the nature of X-bridges". Unfortunately, whilst demanding recognition for our own discoveries we often pay scant regard to other "still small voices" crying in the wilderness of contemporary science. Admittedly authors of undergraduate texts have no choice but to clarify their messages by reducing the surrounding noise of science debate and controversy. Hoyle is no exception in this respect, but I doubt whether dissenters from the vesicle hypothesis for synaptic transmission will appreciate his unquestioning acceptance of this proposal.

In the 1950s it was still possible to pursue a successful career as a comparative physiologist; although it was a struggle to keep abreast of the literature and relevant technological development. It is no longer possible to be a "Jack of all Phyla", as a comparison between Hoyle's original monograph and this new book will clearly show. Despite this, I feel *Muscles and their Neural Control* will be as successful as its predecessor and will, in time, become a classical work. □

P.N.R. Usherwood is Professor and Head of the Department of Zoology, University of Nottingham.

Towards a new phase in physics?

Louis Lyons

Quark Matter Formation and Heavy Ion Collisions.

Edited by M. Jacob and H. Satz.
World Scientific (distributed by Wiley):
1983. Pp. 586. £49, \$70.

PROTONS and neutrons, the constituents of nuclei, are themselves composed of quarks, bound together by gluons. Richard Feynman is quoted as once saying that the chance of discovering anything fundamental in proton-proton collisions is about the same as that of learning about how balance wheels and mainsprings operate by hurling one watch at another. If that is so, what use can there be in attempting to study the collisions of heavy nuclei with each other?

The hope is that, with a large enough number of protons and neutrons (or equivalently, quarks and gluons), a thermodynamic approach may be possible. Furthermore, it is even conceivable that, at high enough energies, the resulting nuclear matter may become so compressed that new phenomena, not accessible in collisions of just a pair of protons, may occur. In particular, increasing attention has been given to the idea that at high densities the quarks and gluons, rather than being confined inside a single proton or neutron, could wander around freely in a quark-gluon plasma extending over the volume of the two nuclei involved in the collision. Such a phase may have existed during the first millisecond of the Universe.

Several crucial questions in this field immediately come to mind. Does such a phase transition in fact exist, and if so under what conditions? Would we expect the new phase to be attainable in heavy ion collisions (and preferably with accelerators already existing)? What would be the best experimental signatures for deducing whether such a quark-gluon plasma had been created?

Such questions were all addressed at length during a workshop held in Bielefeld in May 1982. These published proceedings give a comprehensive view of both experimental and theoretical facets of the subject at the time. They are clearly required reading for nuclear or elementary particle physicists interested in working in this field (although they would have to update their information from material presented at the subsequent conference at Brookhaven in September 1983). The rest of the scientific community, however, can probably safely wait for a few years until some of the basic questions can be answered by the experiments which are now being planned. □

Louis Lyons is a Lecturer in the Department of Nuclear Physics, University of Oxford.

Automatic DNA sequencing

Akiyoshi Wada describes a system that takes the grind out of DNA sequencing

DNA SEQUENCING analysis is of fundamental importance today in pure and applied molecular biology. But the current procedure for analysing DNA sequences is both tedious and time-consuming. In the case of the Maxam-Gilbert scheme, more than thirty steps of microchemical manipulation are involved and these can only be carried out manually by a skilled chemist.

However, even using manual procedures the world-wide rate of analysis may exceed 10^6 bases per year soon, judging by the current rate of additions to the DNA database. But this seemingly high rate of input is far below what is needed to meet the expected demands of sequence analysis in the future. There are many genomes still to be sequenced and even a single human genome has 3×10^9 bases.

An automated DNA sequencer recently developed by Seiko Instruments and Electronics of Tokyo, Japan, should take some of the drudgery out of DNA sequencing and speed up data production. The instrument is a microchemical robot that can carry out the fragmentations of 16 DNA solutions — that is it can test for the four bases (A, G, C and T) in each of four different DNA solutions. It performs the fragmentation in four to five hours, following the Maxam-Gilbert protocol. In

addition, it can carry out two sets of operations sequentially so that 10 hours after the start button is pushed, 32 solutions are prepared for electrophoresis. The analysis may be started in the evening as the laboratory closes and then processed material can be collected the next morning.

The fully automatic chemical operations performed by this Seiko robotic instrument include the quantitative addition of reagents, mixing by vortexing, centrifugal separation, draining, vacuum drying, and the heating and cooling of solutions under treatment. The volume of solutions to be handled can be varied from $1 \mu\text{l}$ to 1 ml. The time schedule and conditions of reaction steps in a series of operations are controlled by a microcomputer command which is input by the operator before the start of the sequencing process. Programmable functions extend this microchemical robot's range of performance to include any of a series of complex chemical reactions of microliter order¹.

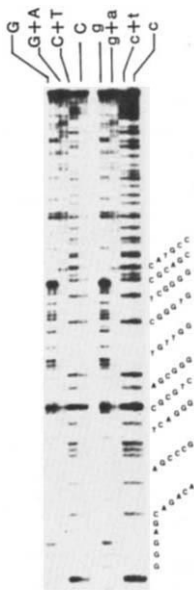
The instrument automates just one aspect of the DNA sequencing procedure and will earn a niche in the market by virtue of its labour-saving capacity rather than by any change in the speed or economy of sequencing. The time it takes to determine a sequence is only a little less than that need-

ed to complete a sequence manually, the sequencer runs the same number of solutions as would a skilled technician and it does not save on reagents. But its designers believe that because it can run for 10 hours when the laboratory is effectively closed and because it frees the chemist to carry out other aspects of the sequencing procedure, it will be commercially viable. Tests have shown that the instrument can perform chemical reactions at the same level of efficiency and reliability as a skilled technician. What is more, the automatic sequencer requires no training and works steadily for 24 hours a day without complaining — something which should make it very attractive to the potential buyer!

The machine should be commercially available later this year, although details of the instrument's price and its exact date of launch have not yet been released.

Akiyoshi Wada is professor of biophysics in the Department of Physics, University of Tokyo and the chairman of the project 'DNA extraction, analysis and synthesis'. The DNA sequencer was developed as part of this project under the auspices of the Science and Technology Council of Japan.

1. Wada, A., Yamamoto, M. and Soeda, E. *Rev. Sci. Instr.* (in the press).



Comparison of sequencing patterns prepared manually (left) and prepared with the sequencer (right).



View of the microchemical robot. The computer system is at the top of the desk and the temperature control dial, stock solutions and pump systems, which are normally under cover, can be seen inside.

General laboratory equipment

● A range of products for use in the DNA laboratory is available from Pharmacia. The products include a centrifugation medium which can be used to separate nucleic acids and proteins. This medium, called CSTFA, allows all types of RNA to be separated by isopycnic centrifugation. Cesium chloride is also available for the preparation of density gradients to purify DNA, and dextran sulphate can be provided to increase the rate of hybridization of nucleic acid molecules. Pharmacia also produces Sephacryl S-1000 which may be used for the removal of low molecular weight material from DNA preparations and in the purification of bacterial plasmid DNA.

Circle No. 100 on Reader Service Card.

● The Schleicher & Schuell Selectasol system is a desk-top chamber for thin layer chromatography (TLC) solvent selection. The Selectasol system is based on the circular development technique. It can be used with TLC ready plates and foils as well as with chromatography papers of between 5×10 cm and 20×20 cm. Up to 16 different solvent systems can be tested on a standard 20×20 cm TLC plate. All the chromatograms develop under the same temperature and relative humidity conditions.

Circle No. 101 on Reader Service Card.

● Two new automatic sample changers are now available from Bioscan. These instruments, the Auto Changer 1000 and 3000, are both compatible with Bioscan's line of radioisotope imaging scanners. Designed to hold from one to three 20-cm thin layer chromatography plates, both instruments provide a system of optical tabs permitting the user to mark up to 60 chromatographic lanes for automated analysis.

Circle No. 102 on Reader Service Card.

● Wheaton Instruments has introduced a new automatic diluter/dispenser which allows samples to be dispensed either manually, semi-automatically or automatically. To accommodate liquids of varying viscosities, the Wheaton diluter/dispenser has an adjustable syringe plunger speed control and an adjustable delay control. Each model is equipped with an interchangeable syringe kit consisting of a sample with a plunger, a reagent syringe with a plunger and a valve assembly. This kit may be removed and autoclaved.

Circle No. 103 on Reader Service Card.

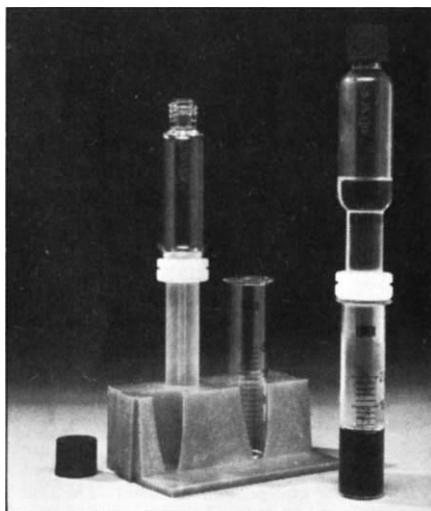
To mark the DNA Winter Symposium taking place in Miami, Florida from 16 to 20 January 1984, this week's selection features products for the molecular biology laboratory. These notes are based on information provided by the manufacturers. For further details circle the appropriate numbers on the Reader Service Card bound inside the journal.

● A new solvent extraction system, designed to replace the separating funnel, is now available from Associated Biolabs. This new Mixxor stirs up the sample and the extraction solvent and then allows them to settle spontaneously into an upper and lower phase. The upper phase can then be decanted without the risk of contaminating it with the lower phase. The Mixxor consists of a glass mixer-reservoir and a glass mixer-separator with a ground glass piston. When fitted together, these can be used in a pumping action. On the 'down' stroke the sample and solvent are forced through a narrow axial channel in the piston into the separator. This produces intimate mixing and leads to very efficient mass transfer. Five strokes of the pumping action produce the same mixing as shaking a separating funnel 40 times. After mixing, the phases are allowed to separate spontaneously. The piston is then adjusted, with the help of a holder-spacer, so that the lower phase is at the very top of the axial channel. The upper phase can then be decanted while the lower phase remains in place. Mixxors are available in four sizes (2, 5, 10 and 20 ml) and are supplied with their own special interlocking racks. These allow an unobscured view of the phase separation.

Circle No. 104 on Reader Service Card.

● An illustrated brochure now available from Beckman Instruments describes the company's system I DNA synthesizer. System I uses the phosphoramidite chemistry, which means that its cycle times are short and side reactions are minimal. The brochure also describes the system I+, which includes the DNA synthesizer and a matched HPLC purification system available as a single package of instruments. Both system I and system I+ can be obtained through Beckman's Base-Pak plan.

Circle No. 105 on Reader Service Card.



The Mixxor from Associated Biolabs

● A new digital syringe diluter/dispenser has been introduced by Dynatech Laboratories. The model 110 is a pipettor/diluter that has a separate dispenser. It has a microprocessor-controlled stepping motor drive and a separate synchronous motor to power the valves. The model 110 can achieve an accuracy of $\pm 1\%$ at total syringe volume. Volumes to be dispensed are programmed in microlitres into the diluter's memory through a 12-key membrane-type keypad which has a four-digit LED display and eight back-lit legends. Stroke rates can also be programmed. Up to 14 programs can be stored in the diluter's memory and are retained even when power is cut off.

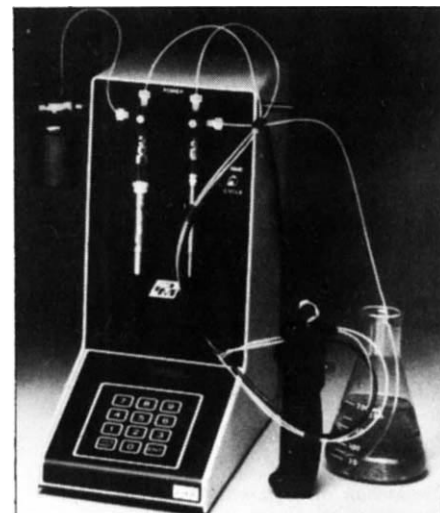
Circle No. 106 on Reader Service Card.

● A line of test tube shakers and rotators has been introduced by Labindustries. The Labquake rotator models provide a constant rotation at 8 r.p.m. Two models are offered which have different capacities and can handle tubes from 10–19 mm diameter to tubes of 27–37 mm diameter, and up to 180 cm long. The Labquake shaker models feature two removable trays. They hold tubes of any diameter up to 162 cm long, and up to 800 g in weight. Speeds may vary from 22 reversals per minute to 70 reversals per minute. Tubes are added and removed while the motor is running. All the shakers and rotators operate on standard 110 V a.c., 50–60 Hz, or 220–240 V a.c., 50–60 Hz. The shakers and rotators are also available in combination models which are supplied with clip bars and interchangeable dual trays.

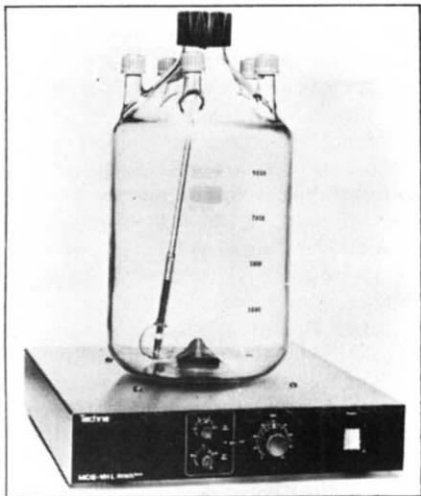
Circle No. 107 on Reader Service Card.

● A line of sample-handling vials and accessories is available from Kontes. These Microflex vials have a conical cavity which makes it easier to retrieve contents with a syringe needle. Five sizes are available — 0.1, 0.3, 1.0, 2.0, 3.0 and 5.0 ml. Optional closures and finishes are available.

Circle No. 108 on Reader Service Card.



Dynatech Laboratories' diluter/dispenser



One of Techne's microcarrier stirrers

● A range of Techne's microcarrier stirrers is now available from Gallenkamp. These stirrers have been designed to increase cell yields by up to 60%. The stirring system is made from borosilicate glass, siliconized in order to inhibit cell growth on the glass surface. A special 'softstart' prevents turbulence on start-up and the subsequent uniform motion of the stirrer rod aids cell growth. Continuous stirring from 0 to 80 r.p.m. or interval stirring from 6 seconds to 5 minutes 'on' and 2 minutes to 2 hours 'off' is possible. The culture vessels are sealed to prevent contamination. The stirrers can be used with an incubator. Four models are included in the range.

Circle No. 109 on Reader Service Card.

● A series of 5cm x 5mm internal diameter mini-columns has been released by HPLC Technology. The columns are packed with the Techsphere, Spherisorb and Hypersil phases and are guaranteed to produce efficiencies of 5,000 theoretical plates per 5 cm column. The 3 μ m Techsphere, Spherisorb and Hypersil ranges include the Silica, C1, C8, C18 (ODS), C18 (ion pair), CN (cyano, NH_2 (amino), phenyl and diol phases).

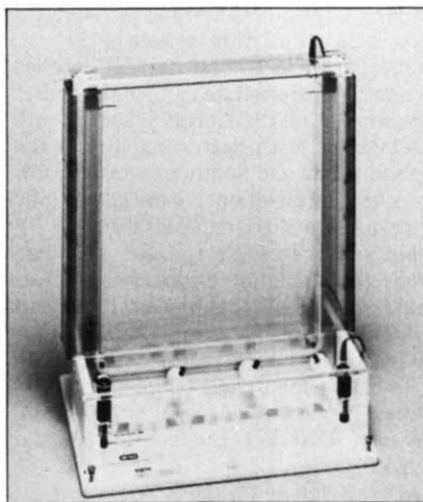
Circle No. 110 on Reader Service Card.

● The AE100 balance has been added to Mettler's AE line of electronic analytical balances. The AE100 has a weighing range of 0 to 109 g. The weighing chamber is accessible from both sides and from the top. A data output option makes it possible to transfer weights to a printer, calculator or computer.

Circle No. 111 on Reader Service Card.

● A range of regulated power supplies from 500 to 4,000 V has been produced by Fotodyne for electrophoresis, DNA sequencing and electrofocusing. The Fotodyne 4,000 V power supply offers constant power, constant current and constant voltage regulation with automatic crossover capability between all modes.

Circle No. 112 on Reader Service Card.



The model 1800 from Bio-Rad

● The new model 1800 DNA sequencing cell from Bio-Rad is designed to reduce the 'smile' effect found when gels are run. The design maintains an even temperature over the gel during electrophoresis. The upper buffer is in contact both with the top of the gel sandwich and with the entire area across one side, so that heat dissipates evenly. Both the upper and lower buffer chambers are incorporated into a single stand that also serves to support the glass sandwich. The sandwich itself is clamped against leakproof seals so that it forms one wall of the extended upper buffer chamber and is in direct contact with the buffer.

Circle No. 113 on Reader Service Card.

● The Drummond Scientific Company has recently added several new dispensing manifold styles to accommodate both 96-well and 24-well microtest plates. For 96-well plates, four-position, eight-position and 12-position manifolds are available. For 24-well microtest plates, 2-position and 6-position manifolds have been added. When attached to the Drummond model 590 digital micro-dispenser, these manifolds can deliver precise aliquots of sample to several positions simultaneously. The manifolds are made of stainless steel.

Circle No. 114 on Reader Service Card.

● A new catalogue detailing BDH calibration kits for gel filtration and gel electrophoresis, SDS polyacrylamide gel electrophoresis and RNA electrophoresis is now available from BDH Chemicals.

Circle No. 115 on Reader Service Card.

● A fluorescence accessory is now available from Beckman Instruments that converts a UV/visible spectrophotometer into a fluorimeter. The accessory can be used with the DU-6 or DU-7 or DU-8B UV/visible spectrophotometer. In using the accessory, the exciting wavelength is first determined by the monochromator setting. The emitted wavelength is then selected with optional filters.

Circle No. 116 on Reader Service Card.

● Tri-Continent Scientific has added an ultra-micropipettor to its line of positive displacement micropipettors. The new pipettor has flexible Teflon tips to facilitate delivery to fragile surfaces, such as those used in electrophoresis, radial immunodiffusion and thin layer chromatography. The pipettor, the Ultra-Micro Absoluter, comes in two models — the 0.5–5 μ l version and the 2–10 μ l version. The capillary tip is reusable.

Circle No. 117 on Reader Service Card.

● Linear Instruments has developed the model 1210 recorder designed for use in liquid or gas chromatography. The model 1210 features a recessed zero adjustment and front-mounted, push button control of power, chart drive and full scale spans. It has an advance/rewind knob for positioning the chart. Chart drive speed is 1 cm min^{-1} , but speeds of 0.25, 0.5, 2 or 4 cm min^{-1} can be achieved through gear changes. The 1210 is a flatbed model which uses 200-mm paper, but it can be wall mounted.

Circle No. 118 on Reader Service Card.

● Heraeus has extended its range of sterile tissue culture vessels by the introduction of Flexiperm. The Flexiperm silicone vessels have an adhesive lower side that sticks to all smooth surfaces to form chambers suitable for culturing cells. The Flexiperm-Slide is used in combination with a standard microscope slide of 75 x 26 mm. The Flexiperm-Disc is intended for use with a Petriperm tissue culture dish, but it may also be placed in a petri dish of polystyrene or glass.

Circle No. 119 on Reader Service Card.

● Wheaton Instruments has introduced the new Step-Pette repetitive syringe dispenser. This delivers pre-set volumes of aliquots from a single draft of fluid. Auto-clavable plastic syringes, supplied with the Step-Pette, permit volumes of 10 to 1,000 μ l to be pipetted each time the thumb knob is depressed. Once the syringe is primed, 10 to 58 dispensings can be performed depending on the volume of the syringe and adjustment setting selected.

Circle No. 120 on Reader Service Card.

ADVERTISEMENT

Liposomes



LIPOPREP® the equipment for the reproducible preparation of stable unilamellar liposomes of defined size in the range of 30-300 nm ϕ with extremely narrow particle size distribution for e.g. the efficient incorporation of (therapeutic) agents incl. antibodies.

DIANORM-Geräte, P.O. Box 126, D-8 München 65, FRG

Circle No. 25 on Reader Service Card.

Biologicals

● *Sau* 1 and *Stu* 1 are the latest introductions to Boehringer Mannheim's programme of restriction enzymes. *Sau* 1 recognizes and cleaves the sequence CC/TAGG producing 5' overlapping ends. It is an isoschizomer of *Cyn* 1 and *Mst* II, and therefore useful in the diagnosis of sickle cell anaemia. *Stu* 1 recognizes and cleaves the sequence AGG/CCT, producing blunt ends. It is an isoschizomer of *Aat* I and *Gdi* I.

Circle No. 121 on Reader Service Card.

● 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside for the visual detection of β galactosidase positive *Escherichia coli* colonies or phage plaques is available from BCL. It comes in pack sizes ranging from 25 mg to 2.5 g. Solutions of X gal at 40 mg ml⁻¹ in *N*, *N'*-dimethylformamide are stable for at least 2 weeks when stored at 4°C.

Circle No. 122 on Reader Service Card.

● New England Nuclear has introduced a new reverse transcriptase system (³²P). This should increase yields of double-stranded cDNA (ds-cDNA) from mRNA templates and eliminate unnecessary steps. The system includes all enzymes necessary to make ds-cDNA as a final product. The enzymes, avian myeloblastosis virus (AMV) reverse transcriptase, *S*₁ nuclease, and DNA polymerase I, are purified and their activities are optimized for the reverse transcriptase system. The system also includes deoxycytidine 5'-triphosphate, (α -³²P)- for labelling reaction products, ribonuclease A inhibitor, standardized buffer solutions and distilled water. All components are supplied ready-for-use.

Circle No. 123 on Reader Service Card.

● Reverse transcriptase that has been purified from avian myeloblastosis virus is now available in the UK from Anglian Biotechnology in conjunction with Dr G.E. Houts. Dr Houts has been providing the product for the National Cancer Institute of America. The transcriptase is allegedly free from contaminating RNase activity.

Circle No. 124 on Reader Service Card.

ADVERTISEMENT

MAP

Analytical Instrument Manufacturers, Custom built spectrometers for geological applications.

For further details see reader reply service.

Circle No. 26 on Reader Service Card.

● New England Nuclear has introduced a new nick translation system (³⁵S) that complements its tritium- and ³²P-labelled systems. This pre-tested system optimizes preparation of DNA probes labelled with dATP α S, (³⁵S) which are suitable for *in situ* hybridization and Southern transfers. For *in situ* hybridizations, autoradiography exposure times are reduced using the ³⁵S label in place of the tritiated analogue, while the resolution necessary for silver grain counting at the light microscope level is maintained. The system is sufficient for 20 reactions and includes deoxyadenosine 5' (α -thio)-triphosphate (³⁵S)-, DNA polymerase I, deoxynucleoside triphosphate mixture, DNase 1 and control plasmid DNA.

Circle No. 127 on Reader Service Card.

● Kor Biochemicals has listed details of its 30 growth factors, cell attachment factors, hormones, transport factors and other related biochemicals.

Circle No. 128 on Reader Service Card.

● Promega Biotec is now manufacturing an *in vitro* translation system. The system is composed principally of New Zealand white rabbit reticulocyte lysate preparation. Other components that are added to the lysate include an enzymatic energy generating system, a mixture of tRNAs and haemin. This means that the two translation-qualified amino acid mixtures, one of which lacks leucine and the other of which lacks methionine, and a control sample of brome mosaic virus mRNA must be added to the solution before use. These items are provided in separate vials. The treated rabbit reticulocyte lysate is provided in 200- μ l aliquots. This eliminates the need for thawing and refreezing. Promega also manufactures and sells RNasin ribonuclease inhibitor, Packagene single extract lambda DNA packing system, GammaPrep labelled nucleotide synthesis systems and restriction endonucleases, including *Sin*I, *Sca*I, and *Dpn*I, as well as DNA and RNA modifying enzymes.

Circle No. 129 on Reader Service Card.

● Bethesda Research Laboratories (BRL) has introduced Nacs Prepac, a prepacked system for nucleic acid purification. Nacs is BRL's family of ion-exchange resins for use in nucleic acid applications and column chromatography. Filled with Nacs 52, a gravity flow resin, Nacs Prepac can be used in plasmid purification and clone screening, removing unincorporated nucleotides from labelling reactions, and recovering nucleic acids from electrophoresis gels. It is available in two forms — the convertible, for use as a mini-column or pipette tip and the cartridge, for use with a syringe. Both the cartridge and convertible are compatible with a variety of common loading and elution accessories. A Prepac contains enough resin to bind up to 40 μ g of A₂₆₀ material.

Circle No. 130 on Reader Service Card.

● Schleicher & Schuell has introduced the S&S Quick-Blot kit. This is a quantitative method for purifying and immobilizing mRNA directly from whole cells or viruses onto nitrocellulose. The new Quick-Blot kit permits selective immobilization of multiple samples of mRNA using dot-blotting techniques in conjunction with the S&S Minifold, a 96-well microsample filtration manifold. The kit can be used to immobilize mRNA from polyribosomes, chromatin, mitochondria, bacteria, yeasts or viruses. Because no baking is required, the mRNA remains biologically active on the nitrocellulose and can be easily translated into polypeptide, reverse-transcribed into cDNA, or released for further manipulation. Each Quick-Blot kit provides sufficient materials to test 120 samples and consists of four basic ingredients: detergents that help rupture the plasma membrane; sodium iodide solutions to promote interaction between mRNA and mRNC and solubilize whole cells at room temperature; NaI-Plus reagent, which furthers the removal of contaminating DNA and protein; and mRNC membrane, an RNA grade nitrocellulose.

Circle No. 131 on Reader Service Card.

● Three new products have been produced by New England Biolabs. The first is a restriction endonuclease, *Hin*P₁ which is derived from *Haemophilus influenzae* P₁. The second is a range of pBR322 *Pst* clockwise site primers for dideoxynucleotide sequencing of DNA. These are specific for the sequence 5' . . . GCTAGAGTAAGTAGTT . . . 3'. The third new product is a series of dephosphorylated pBR322 cloning vectors which includes *Bam*HI cleaved and dephosphorylated pBR322, *Hind* III cleaved and dephosphorylated pBR322, *Eco*R I cleaved and dephosphorylated pBR322 and *Pst* I cleaved and dephosphorylated pBR322.

Circle No. 132 on Reader Service Card.

● Enzo Biochem has added viral specific Patho-Gene kits to the Enzo Bio-Probe system. These kits detect the presence of a virus in fixed cells and tissue preparations. The Patho-Gene kits available contain biotinylated probe DNAs specific for herpes simplex virus I and II, cytomegalovirus, hepatitis B and Epstein Barr virus. The DNA is hybridized and then visualized by a colorimetric enzyme assay for biotin. The Patho-Gene kits are provided with Detek I-hrp, a streptavidin horseradish peroxidase complex.

Circle No. 133 on Reader Service Card.

● Hydroxylapatite, which is used to characterize DNA hybrids and to study reassociation kinetics, is now available from Bio-Rad. It comes in bottles that show the conditions for optimal separation of single and double-stranded DNA on the label.

Circle No. 134 on Reader Service Card.